UNIVERSITY OF SOUTHAMPTON

FACULTY OF MEDICINE

Clinical and Experimental Sciences

Genetic Alterations in Potential Precancerous Skin Conditions

by

Amel Abdulsalam M. Albibas

This thesis is submitted to the University of Southampton for the degree of Doctor of Philosophy

November 2016

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MEDICINE

Clinical and Experimental Sciences

Thesis the degree of Doctor of Philosophy

GENETIC ALTERATIONS IN POTENTIAL PRECANCEROUS SKIN CONDITIONS

Amel Abdulsalam M. Albibas

Non melanoma skin cancer (NMSC) is the most common cancer worldwide. Exposure to ultraviolet radiation (UVR) causes DNA damage in keratinocytes leading to development of NMSCs and precancerous skin conditions such as actinic keratoses (AKs), Bowen's disease (BD) and the clinically invisible p53 mutant patches/p53 immunopositive patches (PIPs). Based on published studies, the rate of progression from these precancerous skin conditions to cutaneous squamous cell carcinoma (cSCC) is variable and the highest rate is seen in AKs (up to 16%) followed by BD (3 - 5%) and the lowest rate of progression is seen in PIPs (1 cSCC for each 8,300 - 40,000 PIPs). The progression from precancerous lesions to cancer is usually accompanied by gain of molecular alterations that affect different genes. Molecular analyses of mutations in lesional DNA, using whole exome sequencing (WES) and targeted enriched sequencing, has demonstrated cSCCs harbour a high mutation burden (33 - 50 mutations/megabase (Mb)). The aim of the current work is to study the mutational status within the genome of precancerous skin conditions, namely AKs, BD and PIPs.

This study was conducted in two stages. In the first stage, AKs were analysed for genetic changes (base pair alterations) using WES. Sections from formalin-fixed paraffin-embedded (FFPE) AKs (N = 69) were initially characterised using histology (for dysplastic area) and immunohistochemistry (p53, beta catenin, CD4, CD8 and FOXP3 protein). DNA from the dysplastic and adjacent normal skin were isolated from 5 AK samples for WES analysis using SureSelect v5 kit (Agilent) and HiSeq 2000 sequencing system (Illumina). WES of the AKs showed that the median mutation burden within AKs was 34.5 mutations/Mb of DNA, with a median number of 1,275 (range 194 – 1,688) mutated genes per AK. In the second stage, 18 genes were prioritised for target enriched sequencing on a wider range of samples including AKs, BDs, cSCCs and PIPs based on the

analysis of the WES/AK results, previous published cSCC WES data and genes mutated in other cancers in Catalogue Of Somatic Mutations In Cancer (COSMIC) database. The samples included 32 additional AKs (including 7 AK/cSCCs where a cSCC was adjoining an AK in the same histological section and considered to have arisen from the adjacent AK) and 37 BDs (including 8 BD/cSCCs where the cSCC was adjoining BD and deemed to have developed from BD), 23 PIPs, as well as corresponding non lesional skin. The target enriched sequencing was conducted using the Truseq custom amplicon kit and MiSeq sequencing system (Illumina), and demonstrated non-silent mutations in PIPs as well as in AKs, BDs and cSCCs. The more commonly mutated genes included *TP53*, *NOTCH1, MLL2* and *HMCN1*. Additionally, the results showed that the genes mutated in PIPs were also mutated in the AKs and BDs, as well as in cSCCs with some identical mutations observed in the different types of lesions. Moreover, analysis of clonality showed some clonal mutations of *CDKN2A*, *CACNA1C*, *MLL2* and *GPR98* genes in PIPs, however, all *TP53* mutations were subclonal. This suggests that in each case the *TP53* mutation occurred after the initial genetic hit which initiated development of the PIP.

Thus, this study has demonstrated that multiple genetic mutations within cancer related genes are present in precancerous skin lesions, including in clinically invisible PIPs that most of them are < 3,000 cells in size, and provides strong support for the “Big Bang” model of cancer in which it is estimated that subclonal evolution begins at an early stage of cancer development.

Table of Contents

List of Tables

List of Figures

DECLARATION OF AUTHORSHIP

I, [Amel Abdulsalam M. Albibas] declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

[Genetic Alterations in Potential Precancerous Skin Conditions]

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:

   Lai, C., August, S., Albibas, A., Behar, R., Cho, S. Y., Polak, M. E., Theaker, J., MacLeod, A. S., French, R. R., Glennie, M. J., Al-Shamkhani, A. & Healy, E. (2016). OX40+ regulatory T cells in cutaneous squamous cell carcinoma suppress effector T-cell responses and associate with metastatic potential. *Clin Cancer Res*, 22,(16), 4236-48.

Signed: ……..

Date: 1/11/2016

Acknowledgements

Where do I start to begin to thank everyone that's helped me over the last 4 years?

Firstly, I would like to thank Professor Eugene Healy for the great supervision. I would like to extend my gratitude to my second supervisor Professor John Holloway for all support and help. Moreover, I would like to thank Professor Sarah Ennis, Dr Matthew Rose-Zerilli and Dr Reuben Pengelly for assistance in exome sequencing data analysis.

I also would like to have a big shout out to my best friend Asma and all my colleagues in the Dermatopharmacology Department, Clinical and Experimental Sciences and Asthma Genetics group, Human Genetics Division, Faculty of Medicine, University of Southampton for their kindness.

All my love and very special thanks for my lovely parents (Kareemah Omar Alhouni Alatshani and Abdulsalam Mukhtar Ammar Albibas) for their support and prayers throughout my life. I would never have been who I am now without them being in my life.

Finally, I would like to thank the Libyan Ministry of Higher Education and Scientific Research for funding this project.

Abbreviations

Abbreviations	Full name
ABC	Avidin-Biotin Complex
AK	Actinic Keratosis
APES	3-Amino-Propyltri-Ethoxy-Silane
ATM	Ataxia Telangiectasia Mutated protein kinase
ATR	Ataxia Telangiectasia and Rad3-related protein
BCC	Basal Cell Carcinoma
BD	Bowen's Disease
BGI	Beijing Genomics Institute
BIRC5	Baculoviral Inhibitor of apoptosis Repeat-Containing 5 gene
bp	base pair
BRAF	Proto-Oncogene, Serine/Threonine kinase gene
BWA	Burrows-Wheeler Aligner
CCDS	Consensus Coding Sequence transcripts
CCND1	Cyclin D1 gene
CDK1-Cyclin B	Cyclin-Dependent Kinase 1 -Cyclin B protein
CDK2-Cyclin E	Cyclin-Dependent Kinase 2 -Cyclin E protein
CDK4	Cyclin-Dependent Kinase 4 protein
CDK6	Cyclin-Dependent Kinase 6 protein
CDKN1A	Cyclin-Dependent Kinase Inhibitor 1A gene
CDKN2A	Cyclin-Dependent kinase Inhibitor 2A gene
cDNA	complementary DNA
cm^2	Square centimetre
c-MYC	Myc Avian Myelocytomatosis Viral gene
COSMIC	Catalogue Of Somatic Mutation In Cancer
cSCC	cutaneous Squamous Cell Carcinoma
DAB	3, 3′ - Di-Amino-Benzidine
DBD	DNA Binding Domain
DNA	DeoxyriboNucleic Acid
ddNTP	2' 3' – dideoxyNucleotide TriPhosphate

dNTP	2’- deoxyriboNucleotide TriPhosphate
DPX	Di-n-butyl-Phalate in Xylene
ECD	Extra-Cellular Domain
EDTA	Ethylene-Diamine-Tetra-Acetic acid
EGFR	Epidermal Growth Factor Receptor protein
FFPE	Formalin Fixed Paraffin Embedded
FISH	Fluorescence In Situ Hybridization
GAPs	GTPase-Activating Proteins
Gb	Gigabase
GDP	Guanosine DiPhosphate
GEFs	Guanine nucleotide Exchange Factors
GPCRs	G Protein Coupled Receptors protein
GPR98	*G protein-coupled receptor 98* gene
GRCh37	Genome Reference Consortium 37
GTP	Guanosine TriPhosphate
H&E	Haematoxylin & Eosin
HES1	Hairy and Enhancer of Split 1 gene
HPV	Human Papilloma Virus
HRAS	Harvey Rat Sarcoma viral oncogene homolog gene
ICD	Intra-Cellular Domain
IHU	ImmunoHistochemistry Unit
IL-10	Inter-Leukin-10
Kb	Kilobase
KIN	Keratinocytic Intraepidermal Neoplasia
KRAS	Kirsten Rat Sarcoma viral oncogene homolog gene
L	Litre
LCM	Laser Capture Microdissection
LNRs	Lin12-Notch Repeats
LOH	Loss Of Heterozygosity
M	Molar
MAML1	MasterMind-Like protein 1
MAPK	Mitogen-Activated Protein Kinase protein

Mb	Megabase
MC1R	MelanoCortin 1 Receptor gene
MDM-2	Murine Double Minutes-2 protein
mRNA	messenger RNA
µg	Microgram
µl	Microlitre
µm	Micrometre
NBF	Neutral Buffered Formalin
NCBI	National Center for Biotechnology Information
NF1	NeuroFibromin 1
NF-KB	Nuclear Factor of Kappa light polypeptide gene enhancer in B-cells gene
ng	nanogram
nm	nanometre
NGS	Next Generation Sequencing
NLSs	Nuclear Localization Signals protein
NMSC	Non Melanoma Skin Cancer
NRARP	NOTCH-Regulated Ankyrin Repeat Protein
NRAS	Neuroblastoma RAS viral (v-ras) oncogene homolog gene
NYAP2	Neuronal tyrosine-phosphorylated Phosphoinositide-3-kinase adaptor 2 gene
OCT	Optimal Cutting Temperature
PARP	Poly ADP-Ribose Polymerase
PCR	Polymerase Chain Reaction
PEST	Proline-, Glutamate-, Serine-, Threonine-rich protein
PI3K	PhosphatIdylinositide 3-Kinases
PIP	P53 Immunopositive Patch
PolyPhen	Polymorphism Phenotyping software
RAF	Root Abundant Factor gene
RalGEF	Ras-like (Ral) Guanine Exchange Factor gene
RAM	RBP-Jκ-Associated Module protein
Rb	Retinoblastoma protein

RBP-Jκ	Recombination signal Binding Protein for immunoglobulin kappa J region protein
RNA	RiboNucleic Acid
SIFT	Sorting Intolerant From Tolerant
SNP	Single Nucleotide Polymorphism
TACE	TNF-α-Converting Enzyme
TAD	Transcriptional Activator Domain
TBE	Tris/Borate/EDTA
TBS	Tris-Buffered Saline
TCGA	The Cancer Genome Atlas
TERT	Telomerase Reverse Transcriptase gene
TIAM1	T-cell lymphoma Invasion And Metastasis 1 gene
Tregs	Regulatory T-cells
TSCA	TruSeq Custom Amplicon Kit
UCSC	University of California, Santa Cruz genome browser
UV	UltraViolet
UVA	UltraViolet A
UVB	UltraViolet B
UVR	UltraViolet Radiation
VAF	Variant Allele Frequency
WES	Whole Exome Sequencing
WHO	World Health Organisation
XP	Xeroderma pigmentosum

Chapter 1: Introduction

1.1 Skin structure

The skin is the largest organ in the human body and represents the barrier between the internal and external environments. It is designed to maintain normal body homeostasis (Wickett and Visscher, 2006). Anatomically, the skin is formed from two main layers; the epidermis and the dermis which are supported upon a layer of subcutaneous fat that plays a role in controlling body temperature (Rees, 2004).

The epidermis is the outer most layer of skin that comes in direct contact with the external environment. The thickness of the epidermis differs across the body (between 50 – 100 μm thick) (Rees, 2004) and is thickest on the palm and soles. The epidermis is highly cellular and formed by a variety of cells with different structures and functions (Haake et al., 2001, Micali et al., 2001). The main cell type is the keratinocyte which are arranged in multiple layers. The deepest layer is known as the basal layer (stratum basale). The basal layer harbours proliferative cells that differentiate while moving toward the surface to form the higher layers, including suprabasal layer (stratum spinosum), granular cell layer (stratum granulosum) and corneal layer (stratum corneum). As they move towards the surface, the keratinocytes lose their nucleus to form the outermost non-nucleated stratum corneum (Micali et al., 2001).

In addition to keratinocytes, the epidermis contains Langerhans cells, melanocytes and Merkel cells (Haake et al., 2001, Rees, 2004). Langerhans cells are migratory dendritic cells that form 3-6% of epidermal cells and represent antigen presenting cells. They are scattered throughout the epidermis to ingest antigenic material such as infectious agents and process them before migrating to the regional lymph node where these cells present the antigen to naive T cells and thus initiate an adaptive immune response (Cumberbatch et al., 2000, Udey et al., 2001). The melanocyte is the epidermal cell which is responsible for the normal variation in colour of the skin (and hair) through the production of melanin. Melanocytes arise from the neural crest (Lin and Fisher, 2007) and migrate to the epidermis where they take up residence amongst basal keratinocytes and within the hair follicle. The number of melanocytes varies at different sites of the skin and the highest frequency is in the genital region (Lin and Fisher, 2007). Racial differences in skin colour are due to the variation in the activity rather than the number of melanocytes (Lin and Fisher, 2007, Rees, 2004). The epidermis also harbours Merkel cells which are restricted

to the basal layer of the epidermis and hair follicle sheath with the highest frequency of Merkel cells within the palmoplantar surfaces (Moll et al., 2005). Merkel cells function as mechanoreceptors due to their synaptic communication with the underneath dermal sensory nerves (Moll et al., 2005). In addition to the various cells, the epidermis contains the upper parts of the skin appendages such as the ducts of sweat glands and the shafts of hair follicles (Haake et al., 2001).

The epidermis is supported and nourished by the underlying dermal layer (~1,000 µm thick) (Rees, 2004). The supportive function of the dermis is due to the high dermal content of connective tissue fibres (collagen and elastin) (Haake et al., 2001, Sorrell and Caplan, 2004). Moreover, the dermis is a highly vascular layer and enriched with lymphatic and nervous supply which are important in providing nutrition and supporting the protective and defence mechanisms of the skin respectively. The cellular component of the dermis involves fibroblast, endothelial cells, and cells which function in skin immunity (T lymphocytes, dendritic cells, macrophages and mast cells) (Haake et al., 2001, Nestle et al., 2009, Sorrell and Caplan, 2004). In addition, the dermal layer is the place of residency for skin appendages which penetrate through the epidermis (hair follicles, sweat glands/ducts).

1.2 Non melanoma skin cancer

1.2.1 Impact and epidemiology

Non melanoma skin cancer (NMSC) is the most common type of human cancer worldwide (Lomas et al., 2012). The two main types of NMSC are basal cell carcinoma (BCC) (accounting for approximately 75% of NMSCs) and cutaneous squamous cell carcinoma (cSCC) (constituting about 25% of NMSCs) (Madan et al., 2010). Both types of NMSC together with their potential precancerous lesions [actinic keratosis (AK) and Bowen's disease (BD)] originate from dysplastic epidermal keratinocytes (Yanofsky et al., 2011). NMSC is a global health problem; in 2006, the World Health Organisation (WHO) stated that the annual occurrence of both cSCC and BCC globally were 2.8 and 10 million cases respectively (Lucas et al., 2006). However, NMSC is more frequent in western societies and the cSCC incidence within Caucasians is around 1 million cases per year (Jemal et al., 2010). The worldwide incidence of NMSC varies across the globe (Lomas et al., 2012). The highest rate is reported in Australia (> 1,000/100,000 person-years for BCC) and the

lowest rate is demonstrated in Africa (< 1/100,000 person-years for BCC) (Lomas et al., 2012). In the USA, more than 5.4 million NMSC cases were treated for NMSC lesions in 2012 (Rogers et al., 2015). The number of registered cases of NMSC within the USA increased from 3.5 million in 2006 (Rogers et al., 2010) to 4.9 million cases in 2011 (Guy et al., 2015). According to Cancer Research UK, the incidence of NMSC in 2013 was 125.2/100,000 population with at least 100,000 cases recorded per annum (Cancer Research UK, 2016).In England, the mean incidence rate is 76.21/100,000 and 22.65/100,000 person-years for BCC and cSCC respectively, with the highest rate in the South West of England (121.29/100,000 and 33.02/100,000 person-years for BCC and cSCC respectively)(Lomas et al., 2012). However, these figures do not represent the actual number of cases as it is recognised that many NMSCs are not recorded (Lomas et al., 2012).

In parallel to the high incidence, NMSC is associated with significant morbidity. Although cSCCs often response well to conventional treatments (surgical excision, radiotherapy), 2.3 - 5.2% of cSCC lesions reappear following these treatments (Rowe et al., 1992). Moreover, cSCCs also have the ability to metastasise and the overall metastasis rate is reported to be 0.3 - 3.7% (Samarasinghe et al., 2011), 3.7% (Schmults et al., 2013), 4% (Brantsch et al., 2008) and 12.5% (Cherpelis et al., 2002) and leads to a 3-year disease-free survival rate of 56% (Martinez et al., 2003) and a 5-year survival rate of 25 - 35% (Kraus et al., 1998, Kwa et al., 1992, Martinez et al., 2003, Rowe et al., 1992). However, larger lesions metastasise to a greater extent with cSCCs > 6 mm deep metastasising in 16% of cases (Brantsch et al., 2008) and with cSCCs > 2cm diameter metastasising in 30 – 40% of cases (Alam and Ratner, 2001).

In the UK, the annual costs of skin cancer management in England was estimated to be £101.6 million (Morris et al., 2009) and is predicted to be over £180 million in 2020 (Vallejo-Torres et al., 2014). In Australia, the annual cost of NMSC alone was around £346 million in 2010 (Fransen et al., 2012). The highest cost was recorded in the USA with the cost of annual skin cancer and precancerous skin lesions management increased from USD$ 3.6 billion in 2002-2006 to USD$ 8.1 billion in 2007-2011 (Guy et al., 2015).

1.2.2 Risk factors of NMSC

NMSC is associated with external (environmental) and internal (genetic) risk factors as listed in table 1.1.

Table 1.1 Risk factors associated with NMSC development. Original table based on information gathered from Berman and Cockerell (2013) Box et al. (2001) Boukamp (2005) De Gruijl and Rebel (2008) Euvrard et al. (2006) Hartevelt et al. (1990) Loeb et al. (2012) Schubert et al. (2014) Tsatsou et al. (2012).

Risk factors	
Exogenous risk factors	**Endogenous risk factors**
1. Chronic UVR exposure (solar or artificial). 2. Ionizing radiation. 3. Cancer causing agents (e.g. arsenic, tobacco smoking). 4. Immunosuppression.	1. Gender (incidence in males > females). 2. Genetic factors (e.g. fair skin due to germline variants in *MC1R* gene and xeroderma pigmentosum due to mutations in genes relevant for nucleotide excision repair).

The main external risk factor for NMSC is accumulative exposure to ultraviolet radiation (UVR) (De Gruijl and Rebel, 2008, Tsatsou et al., 2012). BCC is associated with intermittent exposure to UVR (Kricker et al., 1995), whereas cSCC is related to chronic exposure to UVR (Rosso et al., 1996). In addition, it has been suggested that human papillomavirus (HPV) promotes the transformation of infected keratinocytes to cancer cells through altering the cell cycle and promoting the degradation of p53 protein, but its exact role in skin cancer in the general population is unclear (Michel et al., 2006). Cigarette smoking also increases the risk of cSCC development; in a meta-analysis looking at 25 studies that investigated the relation between smoking and cSCC, a significant relation was found with an odds ratio of 1.52 (95% confidence interval 1.15 - 2.01) (Leonardi-Bee et al., 2012). Immune-status also affects the tendency to develop NMSC, for example organ transplant patients have a 65 (Jensen et al., 1999) to 250 (Hartevelt et al., 1990) fold increased risk of cSCC and a 10 fold increased risk of BCC (Hartevelt et al., 1990). This incidence increases with longer duration of exposure to immunosuppressive agent, for instance, the risk of cSCC increased from 7% after 1 year of immunosuppression to 45 % after 11 years (Bouwes-Bavinck et al., 1996). Moreover, the rate of acquiring a second cSCC lesion increased to 66% in immunosuppressed subjects within 5 years from the appearance of the primary tumour (Tsatsou et al., 2012). The incidence of cSCC also increases in older men and in individuals with fair skin and red hair (Healy et al., 2000, Kennedy et al., 2001).

Germline genetics also play a role in cSCC development. For example, germline variants of the melanocortin 1 receptor (*MC1R*) gene are associated with higher susceptibility to UVR damage and skin cancer (Box et al., 2001, Healy et al., 2000, Healy et al., 2001, Kennedy et al., 2001). The *MC1R* gene is known to influence the colour of skin in a number of vertebrates including humans (Han et al., 2008). The *MC1R* encodes a seven pass transmembrane G-protein coupled receptor that is expressed on the surface of melanocytes and controls the process of melanin formation (Han et al., 2008). Genetic variants in *MC1R* are associated with a decrease in the number of functional MC1R receptors on melanocytes and with reduced signalling via MC1R protein (Beaumont et al., 2012, Rees et al., 1999, Robinson and Healy, 2002), thus altering the total amount of melanin and the phaeomelanin/eumelanin ratio, resulting in fair skin (Healy et al., 2000, Healy et al., 2001, Ito and Wakamatsu, 2003). However, studies have observed that the higher risk of skin cancer due to *MC1R* gene variants remains after correction for skin colour (Bastiaens et al., 2001, Box et al., 2001, Healy et al., 2000, Healy et al., 2001, Kennedy et al., 2001). According to Bohm et al. (2005) *MC1R* gene variants may inhibit DNA repair and reduce apoptosis. In a mouse model, Robinson et al. (2010) found that the number of p53 immunopositive patches (PIPs, i.e. a group of p53 positive epidermal keratinocytes that may develop into cSCC) were lower in UV irradiated *MC1R*+ than in *MC1R*- albino mice. Thus, both the Bohm et al., 2005 and Robinson et al., 2010 studies provide supporting evidence for a non-pigmentary role of *MC1R* gene in photoprotection.

Mutations in certain genes can cause “pathological” pale skin which is associated with an increased risk of skin cancer. The *OCA2* gene mutation is one example which causes type 2 albinism and is common in Africa (e.g. Tanzania), which therefore results in an absence of melanin in skin and reduces protection against UVR, leading to cSCC (Mabula et al., 2012, Yakubu and Mabogunje, 1993). However, genetic variants in a number of other genes have also been associated with a fairer skin type in the normal population and/or an increased risk of skin cancer (although in some cases this association has been with melanoma rather than NMSC). Examples of pigmentation-related genes and some other genes where variants show an association with skin cancer are summarised in table 1.2.

In addition to *MC1R* gene variants, genetic mutations that affect DNA repair mechanisms such as those responsible for xeroderma pigmentosum, increase the susceptibility to UVR

induced mutations and the incidence of skin cancer. Xeroderma pigmentosum is an autosomal recessive disorder that is characterised by a defect in nucleotide excision repair (which repairs UVR induced cyclobutane pyrimidine dimers and 6-4 photoproducts) (Friedberg, 2001) and patients with xeroderma pigmentosum show a >1,000-fold increase in skin cancer risk in comparison to the general population (Schubert et al., 2014).

Table 1.2 Genes with variants associated with increased risk of skin cancer.

Gene	Function
ASIP	Involved in synthesis of phaeomelanin (Lin et al., 2011).
TYR	Involved in the conversion of tyrosine to melanin (Nan et al., 2009).
SLC45A2	It mediates melanin synthesis and is responsible for variations in skin and hair color (Chahal et al., 2016).
OCA2	Involved in small molecule transport, specifically tyrosine. It is involved in mammalian pigmentation controlling skin/eye colors (Chahal et al., 2016, Mabula et al., 2012).
IRF4	Involved in the development and function of T helper cells, regulatory T cells, B cells and dendritic cells and plays a role in development and progression of skin cancer (Chahal et al., 2016).
BNC2	DNA-binding zinc-finger protein thought to act as both a messenger RNA-processing enzyme and a transcription factor14. BNC2 is expressed in melanocytes and, to a lesser extent, keratinocytes, with higher expression levels corresponding to darker skin pigmentation in human skin tissue analysis (Asgari et al., 2016).
CADM1	Metastasis suppressor gene involved in modifying tumors interaction with cell-mediated immunity (Chahal et al., 2016).
AHR	Dioxin receptor involved in anti-apoptotic pathways and melanoma progression (Chahal et al., 2016).
SEC16A	Putative oncogene with roles in secretion and cellular proliferation (Chahal et al., 2016).

1.3 Role of UVR in development of cSCC and precancerous skin lesions

Many studies have shown that exposure to UVR (UVB with wavelengths 290-320 nm and UVA with wavelengths 320-400 nm) is associated with cSCC development (Findlay, 1928, Karagas et al., 2002, Lomas et al., 2012, Milon et al., 2014, Pleasance et al., 2010, Schmitt et al., 2011). One of the earliest studies to demonstrate causation of UVR in skin carcinogenesis was performed by Findlay, 1928, who reported the development of papilloma-like growths on the back of UV irradiated albino mice. The higher incidence of NMSC in Australia is widely considered to be due to proximity to the equator where exposure to UVR is more intense (Lomas et al., 2012). In addition, cSCC is twofold higher in individuals who were exposed to artificial UVR tanning lamps before their twenties (Karagas et al., 2002).

On the other hand, the use of sunscreen reduces the risk of development of cSCC (Green et al., 1999, Stern et al., 1986, Van Der Pols et al., 2006). Stern et al. (1986) reported a reduction in the lifetime incidence of NMSC by 78% as a result of sunscreen application during the first 18 years of life. Moreover, Green et al. (1999) observed a reduction in the rate of cSCC upon daily sunscreen application for 4.5 years and further 40% reduction in

cSCC rate was reported for the entire follow-up duration (8 years) of the same participants (Van Der Pols et al., 2006).

Exposure to UVR can induce skin carcinogenesis mainly by causing DNA damage in skin cells (Brash, 1988, Sage, 1993). A DNA molecule has the ability to absorb UVR which can leads to the formation of DNA photochemical changes in the form of photoproducts (De Gruijl and Rebel, 2008, Svobodova et al., 2006). These photoproducts which comprise cyclobutane pyrimidine dimers (CPDS) and pyrimidine 6-4 pyrimidone photoproducts (6-4PP), as illustrated in Figure 1.1, can result in base pair (bp) changes known as UVR-signature or UVR fingerprint mutations that appear as C > T and CC > TT transitions at dipyrimidine sites (Brash, 1988, Miller, 1985, Patrick, 1977, Sage, 1993, Sinha and Hader, 2002, Wang et al., 2008, Wikonkal and Brash, 1999).

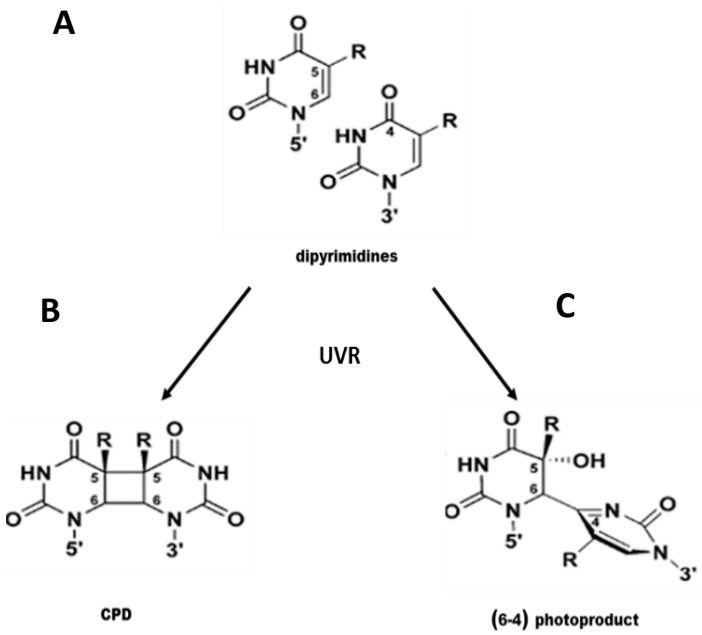


Figure 1.1 UVR induced photoproducts at dipyrimidines in DNA. **A:** Dipyrimidine structure. **B:** cyclobutane pyrimidine dimer (CPD) formation. When the double bonds between C5 and C6 of two adjacent pyrimidine rings are broken, the CPD is generated by the formation of new covalent bonds between the opposite C atoms of the affected adjacent pyrimidines and thus the formation of a cyclobutyl ring. **C:** 6-4 photoproduct (6-4PP) formation. Following the break of the covalent bonds between C5 & C6, a rotation of one of the pyrimidine rings can occur, so that the C6 binds to C4 atom on the adjacent ring (one bond only) resulting in the formation of a pyrimidone/pyrimidine structure called a 6-4 photoproduct (6-4PP). Figure based on information adapted from Wikonkal and Brash (1999).

The pyrimidine dimers cause distortion of the normal alignment of DNA molecules and interfere with the process of replication and transcription (Rastogi et al., 2010). These photoproducts are usually repaired by the nucleotide excision repair mechanism, but if the DNA repair mechanism does not correct these alterations before the end of the S phase of the cell cycle, the DNA of the daughter cells may be mutated because of the substitution of the base pair by a different base pair during the process of DNA synthesis during mitosis (Rastogi et al., 2010). Subsequently, lack of repair of the photoproduct

results in misincorporation of an adenine molecule opposite to the altered cytosine, thus changing the normal guanine to an adenine on the opposite DNA strand which therefore culminates in the substitution of the cytosine by a thymine on the original DNA strand (Matsuda et al., 2000).

Another mechanism by which UVR can promote skin cancer (especially cSCCs) and potential precancerous lesions is through the induction of cutaneous immunosuppression (Kelly et al., 2000, Schwarz, 2005). There is evidence that UVR-induced DNA damage can lead to immunosuppression, for example immunosuppression has been reported in situations of reduced DNA repair (including in xeroderma pigmentosum) (Suzuki et al., 2001) and the administration of exogenous photolyase to assist DNA repair can partially reverse this immunosuppression (Stege et al., 2000). According to Schwarz (2008) UVR induced immunosuppression is mediated by regulatory T-cells (Tregs) that antagonise anti-tumour skin immunity. Tregs, which are FOXP3+CD4+ CD25high T cells, can stimulate immune suppression either by direct inhibition of CD4+ and CD8+ T cells (Fontenot et al., 2005) or through the secretion of interleukin-10 (IL-10) and transforming growth factor beta 1 (Fontenot et al., 2005, Sakaguchi, 2005). In a recent study by Lai et al. (2016) the role of Tregs in cSCC was investigated and functional assays revealed the ability of Tregs within cSCCs to suppress effector T cell proliferation and inhibit interferon-γ secretion by the effector T cells. Another mechanism by which UVR induces immunosuppression is through the production of cis-urocanic acid (Gruner et al., 1992). Urocanic acid is mainly present in stratum corneum in its trans-isomer form. However, upon exposure to UVR, it isomerised to the cis-isomer (cis-urocanic acid). Administration of cis-urocanic acid can mimic the consequences of UVR in the production of transient suppression of contact hypersensitivity responses (Noonan and De Fabo, 1992). It is thought that part of the mechanism responsible for this effect of cis-urocanic acid is through its ability to modulate the production of immune mediators from keratinocytes, nerves and mast cells (Hart et al., 2011, Norval et al., 1990). Another important mediator in UVR induced immunosuppression is IL-10 (Beissert et al., 1996). IL-10 is a cytokine that has a strong effect on skin immunity and tumour immunosurveillance (Beissert et al., 1996) and can suppress the process of antigen presentation both in *vitro* and in *vivo* systems (Fiorentino et al., 1991). In UVB-irradiated skin, IL-10 is predominantly secreted by keratinocytes and

activated macrophages and facilitates down-regulation of contact hypersensitivity responses (Howard and O'Garra, 1992, Pasparakis et al., 2014).

1.4 cSCC pathogenesis

Cancer development is a multi-step process which usually starts with driver mutation(s) (Merlo et al., 2006, Nowell, 1976, Pepper et al., 2009) followed by the accumulation of multiple somatic mutations within the preneoplastic cells (Hanahan and Weinberg, 2011, Klein, 2009, Merlo et al., 2006, Pepper et al., 2009). Many cancers, including cSCC, may be preceded by precancerous conditions. The progression of precancerous lesions to full malignancy is thought to be partially controlled by the tumour microenvironment (Mori et al., 2002, Reid et al., 2010). In the early stages, the "cancer suppressive" mechanisms in the tumour microenvironment can inhibit and/or delay the process of progression, but at a later stage these protective mechanisms may lose their functions (Reid et al., 2010). Furthermore, within the dysplastic precancerous cells, additional somatic mutations in key genes may play a role in the progression of precancerous lesions to cancer. Key genes which are thought to play a role in cSCC development include *TP53*, *NOTCH1* and *NOTCH2* (Bennett et al., 1997, Berg et al., 1996, Boldrini et al., 2003, Bolshakov et al., 2003, Brash et al., 1996, Burns et al., 1993, Caldeira et al., 2004, Caulin et al., 2007, Dumaz et al., 1997, Durinck et al., 2011, Jiang et al., 1999, Kress et al., 1992, Kubo et al., 1994, Lehman et al., 1993, Li et al., 2015, McGregor et al., 1997, Nelson et al., 1994, Oberholzer et al., 2012, Pickering et al., 2014, South et al., 2014, Wang et al., 2011, Ziegler et al., 1994).

1.4.1 Key driver genes in cSCC pathology

1.4.1.1 *TP53* gene and p53 protein

1.4.1.1.1 *TP53* gene

TP53 is a tumour suppressor gene also known as the guardian of the genome (Lane and Crawford, 1979). It is a highly evolutionary conserved gene and is located on chromosome 17p13.1 (Isobe et al., 1986), covering a region of ~20 kilobase (kb) pairs. *TP53* is composed of 11 exons (Lamb and Crawford, 1986) and encodes for the p53 protein that acts as a transcription factor regulating many cellular pathways. The structure of *TP53* gene, transcript and p53 protein are shown in figure 1.2.

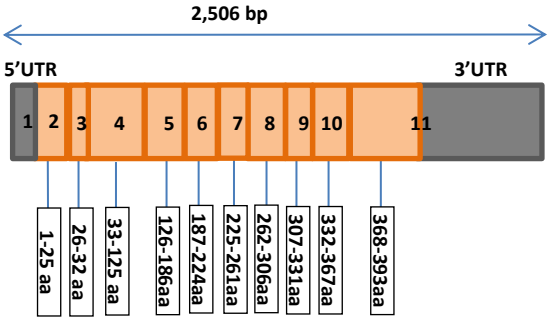
1.4.1.1.2 p53 protein

A. p53 structure

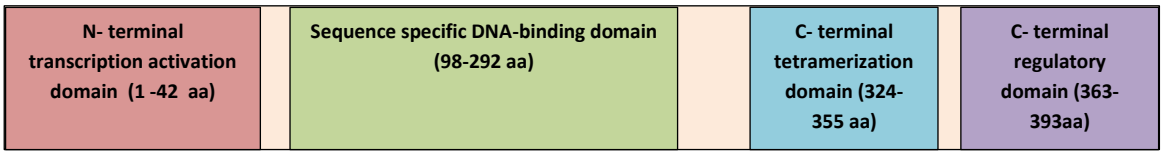
The p53 protein was discovered in 1979 (DeLeo et al., 1979, Lane and Crawford, 1979) and consists of 393 amino acids that are arranged in 4 structural / functional domains. The first 42 amino acids of the protein form the N-terminus which is involved in controlling the transcription of different genes that are important in the regulation of the cell cycle. The next domain is the sequence-specific DNA binding domain or central domain. It covers the region of protein that spans the amino acids residues from 98 to 292. This sequence-specific DNA binding domain is the domain most vulnerable to UVR-induced mutation in keratinocytes (Bode and Dong, 2004). The C-terminus is composed of two domains; a small domain that spans the amino acids residues 324–355 of the protein and is important in protein oligomerisation or tetramerisation, and the last domain (formed of 30 amino acids) known as the regulatory domain and is sensitive to the action of ubiquitin specific protease 7 (Bode and Dong, 2004, Slee et al., 2004, Vousden and Lu, 2002). The regulatory domain has 9 essential amino acids that have DNA and RNA binding activities. p53 protein is usually expressed in limited concentration during homoeostasis (Vousden, 2002). However, its concentration rises in response to cellular stress in order to control the cell cycle checkpoints and to prevent the replication of faulty DNA (Brash, 2006, Vousden, 2002).

A. *TP53* gene

B. *TP53* Transcript



C. p53 protein



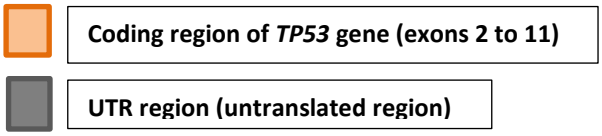


Figure 1.2 Structure of *TP53* gene, transcript and p53 protein. **A:** *TP53* gene; *TP53* exons are shown as rectangles (containing exon number) and introns and non-coding DNA are shown as a blue line. Untranslated exons are depicted in grey, translated exons in orange, with numbers in open rectangles beneath indicating the size of each exon in base pairs. The transcription start (ATG codon) and the end sites are indicated. The position of the gene in the genome is also indicated, along with the length of the *TP53* gene (20 kilobases). **B:** *TP53* transcript; the length of the *TP53* transcript that encodes for the 393 amino acid protein is 2,506 base pairs. *TP53* exons are shown as rectangles (containing exon number) together with the untranslated exons showed in grey and the amino acids of the p53 protein encoded by each exon are indicated in the open rectangles beneath each exon. **C:** p53 protein; the main protein domains are highlighted according to their activity and their amino acid content. aa = amino acid. bp = base pair. Details included in this figure were obtained from Bode and Dong (2004) Lamb and Crawford (1986) (NCBI)(www.ncbi.nlm.nih.gov/gene) and (Ensembl)(www.ensembl.org).

B. p53 expression, regulation and function

p53 is a nuclear protein which is expressed at a high level in response to DNA damage, and through its abilities as transcription factor, it controls cellular pathways associated with cell cycle arrest, apoptosis (programmed cell death), pre-mature cell aging (senescence) and cell differentiation (Amaral et al., 2010, Hamzehloie et al., 2012).

The activity of p53 is regulated by various proteins. These include the E6/E6AP complex whereby the E6 protein from human papillomavirus targets p53 to be degraded (Martinez-Zapien et al., 2016). Another protein which can cleave p53 and reduce the amount of p53 in the cell is calpain, but in some cases, for example following DNA damage, calpain can activate p53 (Sedarous et al., 2003, Zhang et al., 1997). The murine double minute-2 (MDM-2) protein is another regulator of p53 levels. MDM-2 can suppress p53 function by binding directly to the *TP53* gene promoter inhibiting gene expression. Moreover, MDM-2 can act as an E3 ubiquitin ligase and binds through its N-terminus to the p53 protein, thus facilitating the binding of ubiquitin-enzyme E3 to the C-terminus of MDM-2 which results in p53 lysine residue modification and poly-ubiquitination of the p53 protein (Vassilev, 2007). Subsequently, the polyubiquitinated-p53 complex undergoes proteasomal degradation, reducing p53 levels. However, in response to cellular stress, p53 concentration increases within cells. This occurs either by promoting p53 expression as a result of *TP53* promoter acetylation, which lowers MDM-2 binding affinity to the *TP53* promoter, or by inhibition of p53 degradation by phosphorylation-dependant inhibition of MDM-2 activity (Vogelstein et al., 2000, Vousden, 2002). This p53 protein stabilising process is important to ensure the presence of a sufficient amount of functionally active p53 in order to control DNA damage associated with cellular stress (Amaral et al., 2010). However, to maintain homeostasis, activated p53 transcriptionally activates MDM-2 in a process known as an autoregulatory negative feedback mechanism (Wu et al., 1993). Other proteins are also involved in the activation of p53 protein, mainly by phosphorylation or acetylation of different amino acid residues on the protein, resulting in alteration in the conformation of p53. ATM, ATR and PARP proteins are examples of p53 activating proteins which modify p53 via phosphorylation (Lakin and Jackson, 1999, Tong et al., 2001).

During cellular stress, p53 induces cell cycle arrest via increasing expression of the cyclin dependent kinase inhibitor p21 (also known as WAF1/CIP1) and inhibiting cells from entering the S phase of the cell cycle (Latonen and Laiho, 2005). This step is important in preventing damaged DNA from being transmitted to daughter cells and also to allow activation of DNA repair mechanisms (Hamzehloie et al., 2012, Latonen and Laiho, 2005, May and May, 1999). However, when there is extensive DNA damage, apoptosis is triggered, resulting in cell death and therefore preventing survival of a cell with severely damaged DNA (Jeffers et al., 2003). In this case, p53 activates the expression of different pro-apoptotic proteins such as Bax, Apaf1 and Noxa (Jeffers et al., 2003) and simultaneously, p53 represses anti-apoptotic proteins such as Bcl-2 family proteins (Hoffman et al., 2002). In addition, p53 along with other pro-apoptotic proteins is also involved in the activation of the mitochondrial caspase apoptosis cascade during the apoptotic process (Zuckerman et al., 2009).

1.4.1.1.3 *TP53* mutations and cancer

TP53 is inactivated in 50 % of human cancers (Benjamin and Ananthaswamy, 2007, Cho et al., 1994, Hollstein et al., 1991). NMSC and potential precancerous lesions are one of the lesions where *TP53* gene is reported to be mutated (see below). Overall, *TP53* somatic mutations have been identified to be up to 90% within human cSCC (Brash, 2006, Brash et al., 1991).

The early research that highlighted the importance of *TP53* gene in keratinocyte malignancy was reported by Dotto et al. (1988). They noticed the ability of mutant *TP53* to transform papillomas that developed on the back of mice grafted with the p117 keratinocyte cell line (derived in culture from chemically induced mouse papillomas) into carcinoma (Dotto et al., 1988). Based on this finding, two different groups (Brash et al., 1991, Pierceall et al., 1991) sequenced *TP53* from human cSCC lesions and reported *TP53* mutation frequency of 58% and 20% respectively. Moreover, Brash et al. (1991) and Moles et al. (1993) revealed UVR signature mutations as the main base pair changes in mutant *TP53*.

Other groups also reported *TP53* mutations within sporadic human cSCCs (Boldrini et al., 2003, Bolshakov et al., 2003, Brash et al., 1996, Caldeira et al., 2004, Kubo et al., 1994, Nelson et al., 1994, Ziegler et al., 1994), cSCCs in immunocompromised individuals

(Bennett et al., 1997, McGregor et al., 1997, Oberholzer et al., 2012), cSCC cell lines (Burns et al., 1993, Lehman et al., 1993) and cSCC induced in murine models (Berg et al., 1996, Caulin et al., 2007, Dumaz et al., 1997, Jiang et al., 1999, Kress et al., 1992). More recently, exome sequencing studies have demonstrated *TP53* gene mutations in different cSCCs including sporadic cSCC (Durinck et al., 2011, Lee et al., 2014, South et al., 2014), aggressive cSCC (Pickering et al., 2014) and metastatic cSCC (Li et al., 2015).

Most of the human *TP53* mutations within cSCC and AK lesions are located on between exons 4 - 9 with hot spots in codons 173–179, 235–250 and 273–278 (Brash et al., 1991, Ponten et al., 1997, Ziegler et al., 1993). For BD lesions, sequencing of exons 5 -8 of *TP53* gene detected mutations on codons 152-154, 168-171, 192, 247, 248, 250, 254, 265, 273, 278, 279, 280, 284, 286 and 294 of the gene (Campbell et al., 1993b, Takata et al., 1997). Ren et al. (1997) reported that 78% of *TP53* mutations were found to affect the DNA binding domain of the protein and that 91 % of these mutations were missense.

Tp53 mutation has also been investigated in cSCCs in mice (Van Kranen and De Gruijl, 1999). In a study by Van Kranen et al. (1995), *TP53* mutations were detected in 9 out of 30 cSCCs on the back of hairless SKH-HRA mice. The majority of these mutations were located in codon 272, corresponding to codon 278 in human *TP53*.

1.4.1.2 NOTCH pathway

1.4.1.2.1 *NOTCH* genes

The NOTCH pathway is a highly conserved signalling pathway (Blaumueller et al., 1997). In humans, the *NOTCH* gene family encodes 4 receptor proteins (Notch 1, 2, 3 & 4) as listed in table 1.3 (Blaumueller et al., 1997). Each of these receptors is a transmembrane protein with an extra-cellular domain that interacts with 5 different transmembrane-bound ligands encoded by the *DELTA* and *JAGGED* gene families (delta 1, 3 & 4 / jagged 1 & 2 proteins) (Panelos et al., 2008, Sakamoto et al., 2012).

Table1.3 *NOTCH* gene family. The chromosomal location of the genes, number of exons within each gene and the number of amino acids within the resultant proteins are included. Information in the table is adapted from (NCBI) (www.ncbi.nlm.nih.gov/gene) and (Ensembl) (www.ensembl.org/).

Gene ID	Chromosomal localisation	Number of exons in the gene	Number of amino acids in the protein
NOTCH 1	9q34.3	34	141
NOTCH 2	1p12-p11	38	263
NOTCH 3	19p13.2-p13.1	33	2321
NOTCH 4	6p21.3	30	1999

1.4.1.2.2 Notch proteins

A. Protein structure

Notch proteins are hetero-oligomers that span the cell membrane (Brou et al., 2000). The proteins have a large extra-cellular domain (ECD) formed of 29-36 sequential epidermal growth factor (EGF)-like repeats that bind with the N-terminus of the delta/jagged ligand during the process of Notch protein activation (Brou et al., 2000) and three cysteine-rich LIN12-Notch repeats (LNRs) adjacent to the ECD that block nonspecific signalling (figure 1.3). The next domain of the Notch protein is the heterodimerisation domain which immediately precedes the transmembrane component (between the extracellular and intracellular domains) and which mediates heterodimer formation. The transmembrane domain is followed by the intra-cellular domain (ICD) that forms the cytoplasmic portion and is released into the cell cytoplasm upon receptor activation (Sakamoto et al., 2012). The ICD has the following distinct regions: recombination signal binding protein for immunoglobulin kappa J region (RBP-Jκ)-associated module (i.e. RAM) domain, ankyrin repeats, nuclear localisation signals (NLSs) and the proline-, glutamate-, serine-, threonine-rich (PEST) domain. In addition to the general structure of Notch proteins, Notch 1&2 proteins have an extra domain called a transcriptional activator domain (TAD) which is present between the NLS and PEST regions (Amsen et al., 2009, Ranganathan et al., 2011). The *NOTCH* gene structure and activation is illustrated in figure 1.3.

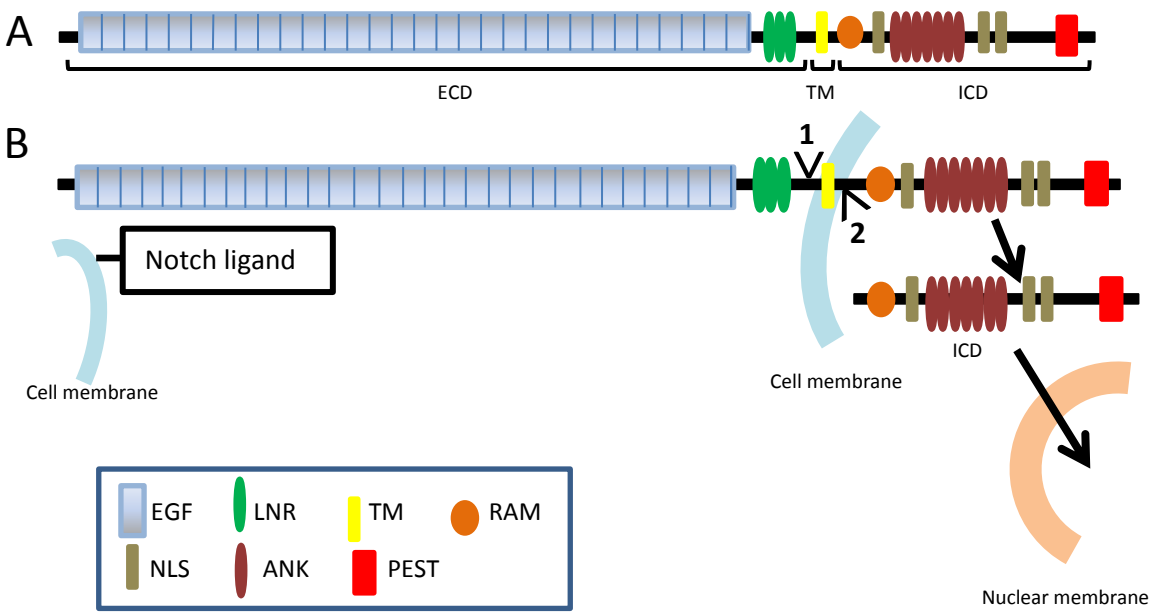


Figure1.3 Notch protein structure and activation. **A:** The general structure of Notch proteins, showing the different domains within the protein. The extracellular domain is formed of epidermal growth factor (EGF) like repeats, three cysteine-rich Lin12-Notch repeats (LNRs) and followed by the transmembrane (TM) part of the protein. The Notch intracellular domain (ICD) is composed of the following: the RBP-Jk-associated module (RAM) domain, nuclear localisation signals (NLSs), ankyrin (ANK) repeats and the proline-, glutamate-, serine-, threonine-rich (PEST) domain. **B:** The process of Notch activation. Interaction of Notch protein with their ligands induces conformational changes within the Notch protein. These changes allow two cleavages to take place (indicated by head arrows 1 and 2) which allows release of the ICD from the cell membrane so that it can enter into the nucleus. Original figure adapted from Ranganathan et al. (2011).

B. Function of notch proteins

Notch proteins have different functions, at both inter- and intra-cellular levels, during different developmental and/ or physiological stages (Capaccione and Pine, 2013). For instance, Notch signalling is crucial in intercellular communication and cell fate decision in order to control healthy growth and differentiation (Capaccione and Pine, 2013). Moreover, it controls the process of auto-regeneration of mature stem cells and the differentiation of progenitor cells in certain cell lineages (Artavanis-Tsakonas et al., 1999). Notch proteins can behave either as a tumour suppressor or tumour promoter during carcinogenesis (Nicolas et al., 2003). For example, Notch proteins can act as oncogenes promoting the growth of malignancy in melanoma (Balint et al., 2005), breast (Ayyanan et al., 2006) and ovarian cancers (Hopfer et al., 2005). In keratinocytes, Notch signalling has an inhibitory effect on the growth of normal (Lefort et al., 2007) and cancerous cells (Wang et al., 2011). Decreased levels of both Notch1 protein and mRNA have been

reported in cSCC cell lines (Lefort et al., 2007) and *NOTCH1* or *NOTCH2* genes mutation frequency is reported in up to75% in cSCC (Durinck et al., 2011, Wang et al., 2011).

1.4.1.2.3 Cascade of Notch activation

The early steps in the canonical pathway of Notch activation is started by receptor–ligand interaction between adjacent cells which stimulates two sequential proteolytic reactions that result in the release of the ICD from the cell membrane which can then move to the nucleus (Amsen et al., 2009, Schroeter et al., 1998). The first proteolytic reaction is mediated by TNF-α-converting enzyme (TACE) and is followed by a second intra-membranous cleavage by the action of γ-secretase resulting in the release of the ICD. Once the ICD has translocated to the nucleus, it forms a complex with RBP-Jk transcription factor and RBP-Jk co-activator (Mastermind-like protein 1 (MAML1)) (Nicolas et al., 2003, Wang et al., 2011). The ICD-RBP-Jk-MAML1 complex binds to the DNA and activates transcription of the *HES1* (HES family BHLH transcription factor 1) gene family (Levitan et al., 2001) thus regulating downstream target genes include cyclin-dependent kinase inhibitor 1A (*CDKN1A)* (Lefort et al., 2007, Nguyen et al., 2006), Cyclin D1 (*CCND1*) (Ronchini and Capobianco, 2001), NOTCH-regulated ankyrin repeat (*NRARP)* (Lamar et al., 2001), nuclear factor of kappa light polypeptide gene enhancer in B-cells *(NF-KB)* (Vilimas et al., 2007), *c-MYC* (Palomero et al., 2006) and baculoviral inhibitor of apoptosis repeat-containing 5 (*BIRC5)* (Chen et al., 2011). In keratinocytes, *CDKN1A* is one of the genes targeted by Notch signalling and promotes terminal differentiation within keratinocytes (Lefort et al., 2007). Moreover, the Notch pathway can crosstalk with the Wnt signal pathway within keratinocytes, which affects beta catenin (a dual action protein that controls cell-cell adhesion and signal transduction) (Brasanac et al., 2005, Nicolas et al., 2003, Ozawa et al., 1989, Papadavid et al., 2002). The interactions of Notch and beta catenin is complex and studies showed that Notch can inhibit beta catenin expression (Acosta et al., 2011). Moreover, activated Wnt/wg signalling increase the cytoplasmic concentration of beta catenin by inhibiting the ubiquitin-proteasomal degradation of the protein (Behrens, 2000). Stabilised cytoplasmic beta catenin then transfers to the nucleus where it binds members of the TCF transcription factor protein family (Behrens, 2000, Huber et al., 1997, Papadavid et al., 2002). This binding activates target genes include *c-MYC* and *Cyclin D1* which subsequently regulate different cellular functions (Xing et al., 2008) including cellular proliferation (Behrens et al., 1998, Liu et al., 2002). Interestingly, a

reduction or loss of the normal membranous beta catenin staining has been reported in 81.8 % of patients with cSCC and/or AKs, with reduced staining more prominent in cSCC than in AK samples (Brasanac et al., 2005).

1.4.1.2.4 Mutation of *NOTCH* genes in cSCC

Nicolas et al. (2003) investigated the role of *Notch* genes in mouse skin epidermis and cornea and found that ablation of *Notch1* is associated with hyperplasia of epidermal and corneal cells and that the absence of *Notch1* also augmented the numbers of chemically-induced skin cancers in this animal model (Nicolas et al., 2003). Moreover, in 2009, Demehri and colleagues provided support for a tumour suppressor role for *Notch1* in murine skin by demonstrating that *Notch1* deleted keratinocytes were prone to papilloma's development (Demehri et al., 2009).

In human cSCCs, *NOTCH* gene family mutations were described by Durinck et al. (2011) and Wang et al. (2011) who reported *NOTCH1 and NOTCH2* mutations in approximately 75% of cSCC samples and cSCC cell line. Recent exome sequencing studies of cSCC (described in section 1.4.1.1.3) also demonstrated *NOTCH1 and NOTCH2* mutations in cSCC (Li et al., 2015, Pickering et al., 2014, South et al., 2014).

1.4.1.3 Other genes mutated in cSCC

In addition to the frequently mutated genes described above, other genes have been seen to be mutated in cSCC and/or precancerous skin lesions with less frequency such as the *RAS* gene family. The *RAS* gene family are group of GTPase genes that are related to the transforming genes of mammalian sarcoma retroviruses (Cox and Der, 2010). They are classified as oncogenes and are involved in signal transduction pathways (Cox and Der, 2010). The *KRAS* (Kirsten rat sarcoma viral oncogene homolog), *NRAS* (neuroblastoma RAS viral (v-ras) oncogene homolog) and *HRAS* (Harvey rat sarcoma viral oncogene homolog) are the original members of a large gene superfamily and were the first human oncogenes identified (more than 30 years ago) (Chang et al., 1982a, Chang et al., 1982b, Cox and Der, 2010, Marshall et al., 1982). The *HRAS, NRAS* and *KRAS* genes share similar structure and function (Malumbres and Pellicer, 1998). The 3 *RAS* genes encode 4 RAS proteins, with 2 KRAS protein isoforms that arise from alternative RNA splicing (KRAS4A and KRAS4B) (Cox and Der, 2010). They have been studied extensively since their discovery and were found to be implicated in different cancers and various physiological

processes that control cell proliferation, differentiation, and survival (Malumbres and Pellicer, 1998) and full details of *RAS* genes cellular function are beyond the purpose of the thesis. In brief, RAS proteins are situated at the inner-plasma membrane of the cells and their activity is under the control of RAS GDP/GTP cycle (Malumbres and Pellicer, 1998, Vigil et al., 2010). The activated RAS-GTP complex is produced upon catalytic reaction involve the transformation of guanosine diphosphate (GDP) to guanosine triphosphate (GTP) with the help of guanine nucleotide exchange factors (GEFs). RAS-GTP complex is deactivated by GTPase-activating proteins (GAPs) that promote GTP hydrolysis (Malumbres and Pellicer, 1998, Vigil et al., 2010). Activated RAS proteins bind to specific downstream effectors (there are at least 11 known discrete RAS effector families, each of which stimulates a different protein signalling pathway of which 3 gene families (*RAF, PI3K* and *RalGEF*) are seen to play a crucial role in the oncogenic RAS pathway) (Kiel et al., 2005, Rodriguez-Viciana et al., 2004, Vigil et al., 2010, Wohlgemuth et al., 2005).

In normal human cells, most of the *RAS* genes are inactive but seem to be activated in 30% of human cancers with missense gain-of-function mutations reported in 27% of human cancers (Bos, 1989, Hobbs et al., 2016, Rodenhuis, 1992, Rodriguez-Viciana et al., 2004). In an agreement with this overall mutation rate of *RAS* genes, further analysis of the mutation frequency of the 3 *RAS* genes using Sanger Catalogue of Somatic Mutations in Cancer (COSMIC) dataset confirms mutation frequency of 22% of *KRAS* within all tumors analysed in comparison to 8% for *NRAS* and 3% for *HRAS* genes (Prior et al., 2012). 98% *RAS* mutations are hotspot mutations that hit a unique position on the RAS protein: G12, G13 and Q61 (Hobbs et al., 2016). Although, *HRAS* is the most extensively studied gene within the *RAS* gene family, 85% of all *RAS* mutations are within *KRAS* followed by 11% in *NRAS*. Interestingly, *HRAS* mutations constitute only 4% of all *RAS* related cancer mutations (Hobbs et al., 2016, Hunter et al., 2015).

RAS mutations negatively affect the GAP-mediated GTP hydrolysis and result in the production of persistently active RAS protein (Hobbs et al., 2016). However, other mechanisms that involve inhibition of GAPs or activation of GEFs are also implicated in RAS activation in cancer (Hobbs et al., 2016). In human cancer, other upstream and downstream components of RAS signalling pathway have also been investigated which includes members of the membrane RTK, cytosolic kinases or RAS/MAPK signalling

pathways and mutations were detected (Fernandez-Medarde and Santos, 2011). For instance, mutations of *BRAF* (proto-oncogene, serine/threonine kinase), one of the downstream targets of the RAS signalling pathway, are detected in 66% of melanoma lesions (Davies et al., 2002).

In skin cancer, *RAS* gene hotspot mutations in codon 12, 13 and 61 have been detected in both melanoma and NMSC. According to the data from COSMIC, the frequency of *HRAS, NRAS* and *KRAS* mutations within human cutaneous BCC is [(7/180), (1/147), (4/147)], within cSCC is [(9/236), (7/107), (5/107)] and within malignant melanoma is [(1/904), (20/3466), (2/924)] respectively (Forbes et al., 2011).

Different studies have been conducted (both in *vivo* and in *vitro*) to investigate *RAS* genes mutations in cSCC and to a lesser extent in precancerous AK and BD lesions. In murine models, *ras* mutations were detected in UVR induced cSCC (Pierceall et al., 1992, Van Kranen et al., 1995) and in chemically (DMBA/TPA) induced cSCC (Lapouge et al., 2011, White et al., 2011). According to Kemp et al., 1993 a combination of *Tp53* gene deletion/mutation and oncogenic *Kras* mutation are necessary to initiate the transformation of the benign DMBA/TPA induced papillomas to cSCC (Kemp et al., 1993). More recently, a whole exome sequencing (WES) study on murine induced cSCC confirmed the results from previous Sanger sequencing based studies and demonstrated *Hras & Kras* mutations in 90% of tested murine cSCC (Nassar et al., 2015).

In human, different exome sequencing studies (Durinck et al., 2011, Lee et al., 2014, Li et al., 2015, Pickering et al., 2014, South et al., 2014) and Sanger sequencing studies (Campbell et al., 1993a, Leis et al., 1998) have demonstrated *RAS* mutations in cSCC.

The role of oncogenic *RAS* in cSCC development in patients on immunosuppressive agents has also been investigated. Oberholzer et al. (2012) showed 21.1% of *RAS* mutations in cSCC developed on BRAF inhibitor treated patients, while Su et al., 2012 demonstrated 60% of cSCC lesions have *HRAS* - Q61L mutation on patients treated with the BRAF inhibitor (Vemurafenib). Paradoxical activation of MAPK (mitogen-activated protein kinase) signalling pathway following BRAF inhibition demonstrated the mechanism by which selective BRAF inhibitors can induce cSCC (Su et al., 2012).

The other gene that is reported to be mutated in cSCC and precancerous skin lesions is *CDKN2A* (cyclin-dependent kinase inhibitor 2A) gene. *CDKN2A* gene is a tumour suppressor gene located on chromosome 9p21 and encodes 2 nuclear phosphor-proteins (p16INK4a and p14ARF) translated from alternative spliced mRNAs. While p16INK4a is translated from exons 1a, 2 and 3 (Kamb et al., 1994, Serrano et al., 1993), the alternative reading frame (exons 1b, 2 and 3) encodes the p14ARF transcript (Mao et al., 1995, Stone et al., 1995). As a consequence, the amino acid sequences of the 2 proteins are distinct (Stott et al., 1998), however, they both have tumour suppressor activity and are involved in cell cycle regulation through the Rb (Retinoblastoma) and p53 signalling pathways (Brown et al., 2004).

The tumour suppressive activity of p16INK4a occurs through the direct and specific binding of the protein to the cyclin-dependent kinases (CDK4 and CDK6) that results in inhibition of kinase depending phosphorylation (inactivation) of Rb (Retinoblastoma) protein. Subsequently, progression of the cell cycle from G0/G1 to S phase is inhibited through the formation of the Rb/E2Fs suppressive complex (Serrano et al., 1993, Sherr and Roberts, 1995, Weinberg, 1995). Cells that lack the p16INK4a activity are likely to gain a proliferative advantage. Thus, p16INK4a is now well known as a tumour suppressor protein in a range of tumours (Sherr, 1996).

p14ARF suppressor function is mediated through its interaction with the E3 ubiquitin-protein ligase MDM-2 that involves in p53 degradation (Quelle et al., 1997). p14ARF overexpression is accompanied by up-regulation of p53 and p21 (Quelle et al., 1997, Zhang et al., 1998). The consequence of this upregulation is either the cessation of the transition of cell cycle from G1 to S phase though the inhibition CDK2-Cyclin E complex (Aleem et al., 2005) or inhibition of the cell cycle at the G2/Mitosis phase through the inhibition CDK1-Cyclin B complex (Dash and El-Deiry, 2005).

Due to the importance of *CDKN2A* in cell cycle control, gene dysfunction through deletion, inactivating mutations, epigenetic silencing or post-translational modification is implicated in carcinogenesis. *CDKN2A* alteration has been demonstrated in a range of cancers including lymphoma (Robaina et al., 2015), pancreatic (Salo-Mullen et al., 2015), lung (Tam et al., 2013), ovarian (Bhagat et al., 2014) and skin (Durinck et al., 2011, Kubo et al., 1994, Lee et al., 2014, Oberholzer et al., 2012, Pacifico et al., 2008, Saridaki et al.,

2003, South et al., 2014) cancers and the alteration status is variable between different cancers (Zhao et al., 2016).

CDKN2A mutations have been investigated in human cSCCs and to a lesser extent in the precancerous AK lesions. The frequency of *CDKN2A* mutations in cSCCs is reported to be between 5.1 % (Oberholzer et al., 2012) and 24.5 % (Pacifico et al., 2008). Brown et al. (2004), described multiple types of *CDKN2A* gene inactivation in human cSCCs. Firstly, they detected loss of heterozygosity (LOH) of 9p21 markers in 32.5% of cSCC cases. Secondly, they detected point mutations in 4 out of 40 tested cSCCs. Lastly, they investigated promoter methylation of *CDKN2A* gene and reported 13 out of 36 cSCCs (36%) had promoter methylation of p16INK4a and 16 of 38 cSCCs (42%) had promoter methylation on the p14ARF alternative spliced isoform (Brown et al., 2004). *CDKN2A* mutations in cSCC have been detected in Sanger sequencing based studies (Küsters-Vandevelde et al., 2010, Oberholzer et al., 2012, Pacifico et al., 2008, Saridaki et al., 2003) and exome sequencing studies (Durinck et al., 2011, Lee et al., 2014, Li et al., 2015, Pickering et al., 2014, South et al., 2014) in cSCC lesions.

1.4.2 Evolution of cSCC

cSCC pathogenesis involves tumour initiation, promotion and progression (D'Orazio et al., 2013, Hanahan and Weinberg, 2011, Seebode et al., 2016). For better understanding of cSCC evolution, the relation between cSCC and its potential precancerous lesions (AK and BD) has been investigated. Precancerous skin lesions show a variable frequency of progression to cSCC (Cox et al., 1999, Criscione et al., 2009, Czarnecki et al., 2002, Fuchs and Marmur, 2007, Jaeger et al., 1999, Kao, 1986, Marks et al., 1988, Mittelbronn et al., 1998). The association between cSCC and precancerous skin lesions is highly supported by the concept of field cancerisation. Field cancerisation is defined as the presence of multiple primary neoplastic or pre-neoplastic lesions that arise within the same area of tissue and was described early in 1953 by Slaughter and colleagues (Slaughter et al., 1953). This area of tissue carries a range of genetic abnormalities induced by chronic exposure to a carcinogen (UVR in case of skin) (Hu et al., 2012, Tsatsou et al., 2012). Thus, characterisation of genetic alterations within precancerous skin lesions may improve understanding of cSCC evolution.

1.5 Potential cSCC precursors

1.5.1 Actinic keratosis

1.5.1.1 Impact and epidemiology of actinic keratosis

Actinic keratosis (AK) is defined as a chronic proliferative epidermal skin lesion that arises from damaged keratinocytes (Ko, 2010, Tsatsou et al., 2012). It is recognised clinically as a scaly or hyperkeratotic lesion with a non-indurated, erythematous base (Ko, 2010, Moy, 2000, Tsatsou et al., 2012). AK is a common global skin condition with increased incidence within certain racial groups and especially in Caucasians (Frost and Green, 1994). AK incidence among adults who are aged ≥ 40 years ranges between 11 - 25% in the northern hemisphere and between 40 – 60 % in Australia (Frost and Green, 1994). In the USA, the incidence of AK in the population aged ≥ 40 years is 60% which cost the health care system USD$900 million annually and is responsible for approximately 8.2 million health care visits per year (Berlin, 2010). In the UK, 24% of people aged over 60 years have at least one AK on sun exposed skin (De Berker et al., 2007).

1.5.1.2 Pathogenesis

Both AKs and cSCC share the risk factors as listed previously in table 1.1 with UVR being the main etiological factor (Ackerman and Mones, 2006, Berman and Cockerell, 2013, De Berker et al., 2007, Frost and Green, 1994, Loeb et al., 2012, Schmitt and Miot, 2012, Tsatsou et al., 2012). Although both AKs and cSCCs arise from uncontrolled division of abnormal epidermal keratinocytes, the dysplastic cells within AK lesions are confined to the epidermis (keratinocytic intraepidermal neoplasia (KIN)) and neither invades the surrounding tissue nor gives rise to metastasis (Bhatia and Spiegelman, 2005, Ko, 2010), thus, disruption of basement membrane integrity is an indicator of the lesion being cSCC rather than AK (Brasanac et al., 2005, Ishida et al., 2001, Ratushny et al., 2012). Moreover, the keratinocytes dysplasia within AK does not involve the full thickness of the epidermis and often occupies the lower two thirds (Bhatia and Spiegelman, 2005, Ko, 2010).

Different studies have defined AK differently; AK has been defined as premalignant lesion (Suchniak et al., 1997), the onset of the SCC continuum (Leffell, 2000), the very early indication of SCC (Lober and Lober, 2000), very superficial SCC (Heaphy and Ackerman, 2000), and neither malignant nor premalignant lesion (Person, 2003).

Despite AK being a common skin lesion, there is a lack of a definitive molecular description for AKs. Thus, improving the classification of AK lesions requires a better understanding of the cellular pathways involved in their development.

1.5.1.3 Histological appearances of AK

AKs show different degrees of keratinocyte dysplasia that has been graded into three main categories; mild, moderate and severe (Brasanac et al., 2005, Cassarino et al., 2006, Cockerell, 2000, Cockerell and Wharton, 2005, Ishida et al., 2001, Quaedvlieg et al., 2006). A number of parameters are used for dysplasia grading which includes; basal cell proliferation, nuclear atypia, mitotic figures, adnexal involvement and dyskeratosis (Boyd et al., 2001). Based on the histological finding, AKs can also be classified into 5 subgroups: hyperkeratotic, acantholytic, bowenoid, atrophic and hypertrophic (Cassarino et al., 2006, Ko, 2010, Yanofsky et al., 2011).

1.5.1.4 Clinical significance of AKs

The importance of diagnosis and management of AK lesions is related to its tendency for malignant transformation (see below). The fact that both AK and cSCCs share the same risk factors, cell of origin and certain genetic alterations (such as those seen in the *TP53* gene) supports the ability of AKs to evolve into cSCC (Bhatia and Spiegelman, 2005, Yanofsky et al., 2011). However, for some patients the importance of AKs is simply because they are sun-induced lesions which look unsightly and for which they would like treatment (Ko, 2010).

1.5.1.5 AK progression rate and time scale

The outcome of a single AK lesion shows wide intra/inter-patient variation. Untreated AKs will results in one of these outcomes: spontaneous regression (Harvey et al., 1996), spontaneous regression with re-appearance (Frost et al., 2000), stable/ persistent lesion or progression to more aggressive disease (Marks et al., 1988). Estimation of AK progression rate to cSCCs is essential to categorise the importance of AK. It has been shown that the risk of AK progression to cSCC is increased with age, number of lesions per patient, a positive history of cSCCs and history of continuous rather than occasional UVR exposure (Alam and Ratner, 2001, Ko, 2010, Rigel and Stein-Gold, 2013). Different figures regarding the AKs progression rate have been identified; Marks et al. (1988) stated that

the annual rate of cSCCs evolution from each AK lesion is 0.075% to 0.096% while Glogau (2000) identified that the annual rate of AK progression to invasive cSCCs is 16%. This rate of progression is similar to an earlier observation by Callen et al. (1997) that showed 20% progression rate per single AK lesion per annum. AKs also have the ability to regress and spontaneous regression is seen in up to 74% of AKs within 12 months of follow up in a study by Frost et al. (2000). Mittelbronn et al. (1998), and Czarnecki et al. (2002) found that 82.4% and 72% of cSCCs were grown on or proximal to actinic keratosis respectively. Moreover, a retrospective studies relying on pathological reports suggested that 60% (Marks et al., 1988) to 80% (Rigel and Stein-Gold, 2013) of cSCC had evolved originally from previously diagnosed AK lesions.

A retrospective study of 91 invasive SCCs that developed on formerly diagnosed AKs (confirmed histologically) using electronic medical records revealed that the mean duration of progression (time scale) to cSCC was 24.6 months (range, 1.97 – 75.6 months) regardless of the gender, age, and location of the lesion (Fuchs and Marmur, 2007).

1.5.1.6 Genes known to be mutated in AK

To date, only few studies have been conducted to investigate the molecular changes on AKs. These include studies that were performed on specific genes that have been shown to be mutated in cSCC lesions. *TP53* mutation frequency within AK ranged between 30 - 60% (Nelson et al., 1994, Nindl et al., 2007, Park et al., 1996, Taguchi et al., 1998, Ziegler et al., 1994). In addition, mutations in *CDKN2A* (Kanellou et al., 2008, Nindl et al., 2007, Pacifico et al., 2008) and *HRAS* (Nindl et al., 2007, Taguchi et al., 1998) have also been reported within AKs. Recently, exome sequencing studies have been conducted on cSCC and revealed a high mutation burden within cSCC (Durinck et al., 2011, Lee et al., 2014, Li et al., 2015, Pickering et al., 2014, South et al., 2014). Thus, the thought was that investigating the genetic profile of AK on a wider range of genes may help in understanding AK pathology.

1.5.2 Bowen's disease

1.5.2.1 Definition and epidemiology

Bowen's disease (BD) or cSCC in situ is defined as full thickness intraepidermal malignancy without sign of invasion to the underlying dermis and recognised clinically as scaly, red plaques that grow overtime on the affected area of skin (Arlette and Trotter, 2004). However, in some cases, these lesions can become hard, warty and/or ulcerated (Arlette and Trotter, 2004). BD usually arises as a solitary lesion, however in 10 - 20% of cases BDs occur at multiple sites (Cox et al., 2007). In parallel with cSCC and AK, BD also originates from dysplastic keratinocytes and affects the elderly (highest incidence in individuals over 60 years) (Eedy and Gavin, 1987, Hansen et al., 2008, Kovacs et al., 1996). BD tends to grow mainly on the ear and scalp (sun exposed areas) in men, and the highest incidence is reported on the head and neck (up to 93.4 %) (Foo et al., 2007, Hansen et al., 2008, Leibovitch et al., 2005). The incidence of BD on the lower legs is greater in women than in men (Eedy and Gavin, 1987, Hansen et al., 2008, Kossard and Rosen, 1992).
A study by Reizner et al. (1994) in Hawaii showed an incidence of 142 of Bowen's disease cases per 100,000 whites. According to the Alberta Cancer Registry, the incidence of BD cases form between 10 - 15 % of all NMSC cases reported between 1988 and 2007 and thus, over 350,000 cases were expected per year in USA (Jung et al., 2010). In the UK, the British Association of Dermatologists' guidelines for the management of squamous cell carcinoma in situ (BD) 2014 (Morton et al., 2014) estimated the annual incidence of BD to be 15 cases per 100,000 population (an analysis based on USA data by Chute et al. (1991)).

1.5.2.2 Risk factors and histological changes

As with cSCC and AK, BD shares the same risk factors with cSCC and AK (as listed previously in table 1.1) with UVR being the main etiological factor (Kossard and Rosen, 1992, Reizner et al., 1994). The risk factors involved in BD pathogenesis include exposure to arsenic (Yu et al., 2006), Immunosuppressive agents (Drake and Walling, 2008) and human papilloma virus (HPV) infection (Hama et al., 2006, Murao et al., 2014).

The degree and the extent of keratinocyte dysplasia differ between the cSCC, AK and BD lesions and being greater in cSCC than in BD and AK (Lohmann and Solomon, 2001). The dysplasia is full thickness in BD (Christensen et al., 2016) and the cellular atypia is frequently extended to involve adnexal structures (Cassarino et al., 2006).

1.5.2.3 The risk of BD progression/regression

BD has the tendency to progress to cSCC but less than that of AK (Morton et al., 2014) which may be due to the limited number of studies on BD in comparison to that on AK. The rate of BD progression was estimated to range between 3 - 5 % (Kao, 1986). However, other observations showed ~ 30 – 50 % of BD patients may have previous or subsequent NMSC lesions (Reizner et al., 1994). The standardised incidence ratio (the ratio of observed-to-expected numbers of cancers) of NMSC is 4.3 in patients with BD (Jaeger et al., 1999).

A limited number of studies have shown a tendency of BD to regress. Murata et al. (1996) found a partial regression of BD in a group of 17 patients. Moreover, Nihei et al. (2004) described one BD Japanese case where the lesions (4x3 cm in size) disappeared spontaneously 2 years after diagnosis.

1.5.2.4 Alteration in molecular profile and protein expression within BD

The number of genetic studies on BD is limited. *TP53* gene mutations were detected in 8 out of 20 BD cases (Campbell et al., 1993b) and in 9 out of 29 BD cases (Takata et al., 1997).

The expression of different proteins demonstrated to be altered in cSCC has also been investigated in BD lesions (Bagazgoitia et al., 2010, Sakiz et al., 2009, Talghini et al., 2009). Relying on immunostaining, the rate of p53 expression in BD was 41.8% (Talghini et al., 2009) and 100% (Bagazgoitia et al., 2010). Bagazgoitia et al. (2010) has also investigated the expression of p16INK4a protein within BD samples (N = 5) and found strong cytoplasmic and nuclear positivity in all samples. In addition, the expression of Ki67 protein (proliferation marker protein that shows nuclear expression during interphase of the cell cycle) (Scholzen and Gerdes, 2000) was 23.7, 12.3 and 19.3% within AKs, BDs and cSCCs, respectively (Talghini et al., 2009).

1.5.3 Epidermal p53 mutant patch/ p53 immunopositive patch (PIP)

1.5.3.1 Background and pathogenesis

Chronic sun exposed normal looking skin harbours clusters of p53 immunopositive epidermal cells called p53 mutant patches / p53 immunopositive patches (PIPs) (Berg et al., 1996, Brash et al., 1996, Jonason et al., 1996, Kanjilal et al., 1995, Rehman et al., 1994,

Rehman et al., 1996, Ren et al., 1997, Tabata et al., 1999, Urano et al., 1995, Ziegler et al., 1994). These PIPs are clinically invisible and can be visualised only after immunostaining of whole mount epidermal sheets using anti-p53 antibodies. The PIPs represent a clonal expansion of p53 immunopositive keratinocytes, many of which harbour *TP53* mutations (Brash et al., 1991, Jonason et al., 1996). p53 mutant keratinocytes are usually apoptosis resistant and gain a growth advantage over normal epidermal keratinocytes (Jonason et al., 1996, Ziegler et al., 1994). Each clone is thought to be formed from the proliferation and subsequent expansion of a mutant progenitor keratinocyte that sustained DNA damage in the *TP53* gene which was not repaired (Zhang et al., 2001b). This model of clone formation is supported by the presence of similar *TP53* gene mutations from cells isolated within the same clone (Zhang et al., 2001b). However, the presence of different *TP53* gene mutations in one clone can be explained by increased genetic instability in the mutant cells that make them prone to more mutations (Zhang et al., 2001b). PIPs are thought to be originated from epidermal basal stem cells and can be detected within interfollicular epidermal-dermal junction as well as around hair follicles (Jonason et al., 1996, Ziegler et al., 1994).

The number and the size of PIPs are largely related to exposure to UVR. Hence, within sun exposed human skin the number of PIPs can reach 33 patches/cm², in comparison to an average of 3 patches/cm² within sun shielded skin (Jonason et al., 1996). The size of a PIP differs widely from 60-3000 cells/cluster within sun exposed skin (Jonason et al., 1996). Moreover, the number of PIPs increases with age and the density of PIPs has been reported to be higher in the skin surrounding NMSC lesions which is in favour of a role of PIPs in NMSC pathogenesis (Backvall et al., 2004b, Ren et al., 1996).

1.5.3.2 The outcome and clinical significance of PIPs

PIPs are thought to persist. In a murine model, UVR-induced PIPs that had disappeared upon cession of UVR treatment, re-appeared after single dose of UVR (Lu et al., 2005). Persisting PIPs are thought to be associated with an increased risk of NMSC development (Roshan and Jones, 2012). The selective growth advantage of certain PIP mutations has been debated. Facts such as both PIPs and cSCCs exhibiting a dose-time dependency on UVR in a mouse model (Rebel et al., 2001), both PIPs and cSCCs sharing UVR-signature hotspot mutations (C > T and CC > TT) (Jonason et al., 1996), the rise in the percentage of

these patches in Xeroderma pigmentosum (XP) patients (who have a defect in DNA repair) (Williams et al., 1998), and in UVR-exposed Xpa-deficient mice (Rebel et al., 2001) support a connection between PIPs and cSCCs. On the other hand, the observation of a low incidence of cSCCs compared with PIPs (~ 1 cSCC for 10,000 PIPs) has led some researchers to question this association. The chance that a PIP progresses to cSCC is relatively small (estimated at 1 out of 8,300-40,000/individual when the first tumour appears), (Rebel et al., 2005).

1.5.3.3 *TP53* gene mutations within PIPs

According to (Backvall et al., 2004a), around 60% of PIPs carry *TP53* gene mutations of which 78% are UVR signature mutations. Analysis of the *TP53* mutation spectra within PIPs from human epidermis adjacent to cSCCs revealed that all mutations were point mutations (missense mutations) and most of them were located in the conserved domain of p53 (Backvall et al., 2004a). The mutations were distributed over exons 4 - 8 of *TP53* gene and hit the following codons; 119, 157, 160, 161, 173, 177, 178, 180, 197, 205, 235, 236, 245, 247, 248, 249, 250, 255, 278, 279, 286 (Backvall et al., 2004a). Moreover, LOH within *TP53* gene have also been investigated. According to Rehman et al. (1996), 64% of AKs showed LOH within the short arm (p) of chromosome 17 that encodes *TP53* gene and this percentage rose up to 80% in the study by Ahmadian et al. (1998). In contrast, another study showed there was no evidence of LOH in *TP53* within the DNA extracted from PIPs (Backvall et al., 2004a). Compound homozygous mutations that hit both *TP53* alleles have been observed in PIPs (Kramata et al., 2005). Despite all these studies, association between cSCC and PIPs is not fully understood. Thus, expanding the range of genetic studies on PIPs might give an insight into PIPs as precancerous skin conditions.

1.6 Hypothesis

Cutaneous squamous cell carcinoma (cSCC) shares mutations on key genes with the potential precancerous lesions (actinic keratosis (AK) and Bowen's disease (BD)), and PIPs.

1.7 Aims of this thesis

The aims of this thesis are;

1. Establish methodologies for the extraction of DNA with suitable quality and sufficient quantity for exome sequencing analysis from formalin fixed paraffin embedded (FFPE) skin biopsies and from PIPs in skin epidermis.

2. Define genetic changes (base pair changes) of DNA samples from AK lesions using whole exome sequencing.

3. Target enriched sequencing on a wider range of potential precancerous skin lesions including AK and BD samples.

4. Target enriched sequencing on DNA samples extracted from the PIPs detected within whole mount normally looking skin epidermis samples lying adjacent to NMSC lesions.

Chapter 2: Materials and Methods

2.1 Materials

2.1.1 Consumables, kits and instruments

Laboratory consumables, kits and instruments for the work conducted in this thesis were obtained from a variety of sources, as listed in table 2.1. All reagents were stored at room temperature, unless stated otherwise.

Table 2.1 Table of consumables, kits and instruments.

A. Table of consumables and their suppliers.

Consumables	Suppliers
Immunohistochemistry and tissue fixation	
3-aminopropyltriethoxysilane (APES) slides	Sigma-Aldrich, UK
Xylene	Fisher Scientific, UK
Ethanol	Fisher Scientific, UK
Hydrogen peroxide (H_2O_2) (stored at 4^oC)	Sigma-Aldrich, UK
Methanol	Fisher Scientific, UK
Citric acid crystal ($C_6H_8O_7$)	BDH laboratory supplier, UK
Sodium hydroxide (NaOH)	BDH laboratory supplier, UK
Sodium chloride (NaCl)	Fisher Scientific, UK
Hydrochloric acid (HCl)	Fisher Scientific, UK
EDTA (Ethylene-Diamine-Tetra-Acetic acid)	BDH laboratory supplier, UK
Avidin/Biotin blocking kit (stored at 4^oC)	Vector laboratories, UK
Vectastain elite ABC kit (standard) (stored at 4^oC)	Vector laboratories, UK
Rabbit anti- human CD4 protein monoclonal antibody - clone [EPR6855] (stored at 4^oC)	Abcam, UK
Mouse anti- human CD8 protein monoclonal antibody - clone [C6B5F5] (stored at 4^oC)	Abcam, UK
Mouse anti- human FOXP3 protein monoclonal antibody - clone [236A/E7] (stored at 4^oC)	Abcam, UK
Mouse anti- human beta-catenin protein monoclonal antibody - clone [E-5] (stored at 4^oC)	Insight Biotechnology Ltd, UK
Monoclonal mouse anti-human p53 protein-clone DO-7 (stored at 4^oC)	DAKO, UK
Polyclonal rabbit anti-mouse immunoglobulin biotinylated (stored at 4^oC)	DAKO, UK
Polyclonal swine anti-rabbit immunoglobulin biotinylated (stored at 4^oC)	DAKO, UK
Foetal bovine serum (FBS) (stored at 4^oC)	Fisher Scientific, Gibco, UK
Bovine serum albumin (BSA) (stored at 4^oC)	Sigma-Aldrich, UK
Dulbeccos modified eagle medium (DMEM) (stored at 4^oC)	Fisher Scientific, Gibco, UK
DAB substrate chromogen system (stored at 4^oC)	DAKO, UK
Liquid DAB substrate pack (stored at 4^oC)	BioGenex, UK
Mayers haematoxylin	Sigma-Aldrich, UK
DPX mounting media	Fisher Scientific, UK
Cover slips	BDH laboratory supplier, UK
Dumont 3c forceps standard tip	Fine Science Tools
Surgical scalpel blade No.14	Swann Morton
Surgipath-neutral buffered formalin concentrate	Leica Biosystem, UK
Saponin	Sigma-Aldrich, UK
TRIS	Fisher Scientific, UK
Slide tray 82mm x 156mm 24 slide capacity	Fisher Scientific, UK
Vectashield® mounting medium (stored at 4^oC)	Vector laboratories, UK

Consumables	Suppliers
Haematoxylin and Eosin (H&E)	
Mayers haematoxylin	Sigma-Aldrich, UK
Eosin yellow	TCS Biosciences Ltd, UK
Ethanol	Fisher Scientific, UK
Xylene	Fisher Scientific, UK
Laser Capture Microdissection (LCM)	
Cresyl violet acetate ($C_{18}H_{15}N_3O_3$)	Sigma-Aldrich, UK
Nuclease free water	Invitrogen, UK
LCM pen membrane glass slides	Life Technologies Ltd, UK
Straight dissecting needle No. 72948	Thomas Scientific, UK
ALT tissue lysis buffer	Qiagen, UK
Nucleaseliminator™	AMRESCO, UK
Molecular experiments	
10X TBE buffer (Tris-Borate-EDTA)	Fisher Scientific, UK
Nancy-520	Sigma-Aldrich, UK
MyTaq™ Red DNA Polymerase kit (stored at -20^{o}C)	Bioline Reagents Ltd, UK
Customise DNA oligos (*KNSTRN* exon-1) (stored at -20^{o}C)	Eurofins MWG Operon, UK
Customise DNA oligos (*MC1R*) (stored at -20^{o}C)	Eurofins MWG Operon, UK
Customise DNA oligos (panel / multiplex PCR) (stored at -20^{o}C)	Eurofins MWG Operon, UK
HyperLadder™ 100bp (stored at 4^{o}C)	Bioline Reagents Ltd, UK
Agarose	Sigma-Aldrich, UK

B. Table of Kits and systems used for sequences.

System	Supplier
QIAamp DNA FFPE tissue kit	Qiagen, UK
SureSelect v5 kit	Agilent technologies, UK
TruSeq custom amplicon v1.5 kit	Illumina, UK
TruSeq PE cluster v3 Kit	Illumina, UK
MiSeq system sequencing	Illumina, UK
HiSeq system sequencing	Illumina, UK
Applied Biosystems 3730/3730xl system sequencing	Thermoscientific, UK

C. Table of instruments.

Instrument	Supplier
Nikon Eclipse 80i microscope	Nikon Instruments, UK
S10 Stage Counting Grid, 0.05mm²	Pyser-SGI, UK
AS LMD laser microdissection microscope	Leica microsystem, UK
GXM-XTL3101, 7X - 45X, stereo zoom microscope	GT Vision Ltd, UK
EMD Millipore Synergy™ ultrapure water purification system	Fisher Scientific, UK
Wesbart IS89 multi shelf shaker incubator	Bidspotter, UK
Nanodrop D-1000	Nanodrop Technologies, US
Thermocycler (GeneAmp ®, PCR System 9700)	Applied Biosystems, UK
Microwave	Sharp compact, UK
RM2235 rotary microtome	Leica microsystem, UK
Leica Biosystems CM3050 S research cryostat	Leica microsystem, UK
DigiDoc-It® imaging system (UVB)	LLC. Upland, CA, US
Qubit® dsDNA BR assay kit	Invitrogen, UK

2.1.2 Buffer and reagent preparation

2.1.2.1 0.01 M Citrate buffer

2.1 grams of citric acid crystal ($C_6H_8O_7$) was dissolved in 1 litre (L) of deionised water and the pH adjusted to 6 using 1 molar (M) of sodium hydroxide (NaOH); the NaOH was added as drops using a Pasteur pipette to the citric acid solution stirring on a magnetic stirrer.

2.1.2.2 20 mM EDTA

A stock solution of 0.5 M EDTA (Ethylene-Diamine-Tetra-Acetic acid) was made by dissolving 186.1 grams of EDTA-$Na_2.2H_2O$ into 800 ml of deionised water. While stirring on a magnetic stirrer, the pH was adjusted to 8 using 1M NaOH and the 0.5 M EDTA stock solution then stored at room temperature in clean glass bottle. Subsequently, 20 mM EDTA was freshly made upon requirement by adding 1 ml of the 0.5M EDTA stock solution to 24 ml of Tris buffered saline (TBS).

2.1.2.3 Tris buffer saline (TBS)

A mixture of 80 grams of sodium chloride (NaCl) and 6.05 grams of Tris (i.e. Tris(hydroxymethyl)aminomethane, $(HOCH_2)_3CNH_2$) was added to 1 L of deionised water. The pH of the solution was adjusted to 7.6 by adding 1 M hydrochloric acid (HCl) and then another 9 litres of deionised water was added to make a final 10 L TBS solution.

2.1.2.4 Blocking media

Blocking media for immunostaining of FFPE sections was made by mixing 1% bovine serum albumin (BSA) and 20% foetal bovine serum (FBS) in Dulbeccos Modified Eagle Medium (DMEM). Blocking medium of 3% BSA and 15 % FBS in DMEM was used for p53 immunostaining of epidermal sheets.

2.1.2.5 3.3′ - di-amino-benzidine (DAB) solution

DAB solution was prepared according to the manufacturer's protocol by adding 1 drop of DAB chromogen to 1 ml of the substrate buffer containing H_2O_2 (DAKO, UK) or it was prepared by the addition of 2 drops of the DAB solution and 1 drop of H_2O_2 to 250 µl of substrate buffer and 2.25 ml of deionised water (BioGenex, UK).

2.1.2.6 1% Cresyl violet acetate solution

1% w/v of cresyl violet acetate solution was prepared by dissolving 1 gram of cresyl violet acetate ($C_{18}H_{15}N_3O_3$) within 100 ml of 100% ethanol at room temperature through stirring on a magnetic stirrer.

2.2 Methods

For molecular techniques, all reactions were performed within autoclaved disposable (used only once) plastic ware and all reactions were prepared using nuclease free water. All services were cleaned with 70% ethanol and nucleaseliminator (AMRESCO). To reduce the risk of inter-sample contamination, filtered tips were used.

2.2.1 Source of human tissue

2.2.1.1 Archived FFPE blocks

The University Hospital Southampton NHS Foundation Trust electronic case records system was used to identify archive FFPE blocks from 113 patients with different skin conditions; AK (N = 69), BD (N = 29) and cSCC adjacent either to AK or BD lesions (N = 15). The search process involved using word searches on the Histopathology results database to identify relevant samples; phrases used as word searches included ‘actinic keratosis and skin’, ‘Bowen’s disease and skin’ and ‘cutaneous squamous cell carcinoma and actinic keratosis or Bowen’s disease’. The Histopathology report for each sample was then checked using the University Hospital Southampton NHS Foundation Trust electronic records system. Samples with histological reports other than the targeted lesions (AK, BD and cSCC) and punch biopsies were omitted from the selection. FFPE blocks containing AK, BD and cSCC adjacent to either AK or BD lesions from the above search were subsequently collected from the Histopathology Department, University Hospitals Southampton NHS Foundation Trust. Ethical approval for this study was obtained from the local research ethics committee (LREC number 07/H0504/187) and a signed consent form indicating that the tissue could be used for research following its initial histopathological use for diagnostic purposes was obtained from the records of all subjects.

2.2.1.2 Redundant breast skin

Redundant normal looking breast skin was obtained from patients who underwent mastectomy for breast cancer (5 females; aged between 36 - 54 years). The surgery was performed by the breast surgery team on patients with confirmed/suspected breast cancer in Princess Anne Hospital/Southampton/UK. Ethical approval for this investigation was obtained from the local ethics research committee (LREC number 07/Q1704/59) and a signed consent form indicating that this surplus tissue could be used for research was obtained from the records of all subjects.

2.2.1.3 Blood samples

Blood DNA samples were obtained from patients as part of an investigation on skin cancer (LREC number 07/H0504/187) being conducted by Dr Chester Lai in Dermatopharmacology, University of Southampton. A signed consent form was obtained from all subjects for the use of the tissue / blood in the research.

2.2.1.4 Fresh normal looking skin adjacent to suspected cSCC lesions

Normal looking skin samples were collected from redundant skin taken to enable closure of the wound following excision of skin lesions; this skin was located at least 4 mm away from lesion being excised (which included cSCC and/or BCC lesions). Most of the normal looking skin samples were obtained from sun exposed skin areas. Skin samples from a total of 126 patients were obtained, under local research ethics committee approval (LREC number 07/H0504/187) from Dr. Chester Lai, Dermatology Department, University Hospital Southampton NHS Foundation Trust. A signed consent form was obtained from all subjects.

2.2.2 Section/slide preparation

5 µm tissue sections (for H&E and immunostaining) and 10 µm tissue sections (for LCM and DNA analysis) were cut from FFPE blocks using a Leica RM2235 rotary microtome. Tissue sections were then placed onto slides pre-treated with 3-aminopropyltriethoxysilane (APES) (for H&E and immunostaining) or onto LCM pen membrane glass slides (for LCM and DNA analysis). Sections were left to dry in incubator at 37 ˚C (may take several hours). For the LCM work, all equipment was cleaned carefully

with 70% ethanol and nucleaseliminator decontamination wipes and nuclease free water (EMD Millipore Synergy™ ultrapure water purification system) was used as required.

To look at p53 immunopositivity in vertical sections of the epidermal sheets that were positive for PIPs, the epidermal sheets were mounted vertically within a drop of OCT (Optimal Cutting Temperature) and allowed to freeze within the cryostat chamber at -20 ˚C. Then 5 µm tissue sections were cut using a cryostat microtome (CM3050 S, Leica Biosystems) for assessing the depth of p53 staining within the epidermal layers and for subsequent counterstaining with haematoxylin.

2.2.3 Separation and fixation of epidermal sheets

Following the receipt of normal looking skin, the subcutaneous adipose tissue was scraped off using a disposable scalpel (Swann Morton) and then the skin samples were divided into small pieces of ~ 2 x 3 mm^2. The epidermal sheet was separated by immersing the tissue in 20 mM EDTA in TBS in microcentrifuge tubes at 55 °C for 5 minutes. Then the tissue pieces were transferred immediately into the wells of a sterile 12 well plate containing TBS at 4 °C and left to cool for 5 minutes (to stop the effect of heat on the tissue). The skin was subsequently placed (with epidermis side up) on a clean petri dish, and the epidermal sheets gently separated from the dermis manually by pulling the epidermis and dermis apart using two fine smooth curved forceps. The epidermal sheets were then fixed by overnight incubation with 10% neutral buffered formalin (NBF) (Leica Biosystem) at room temperature and stored (for a maximum of 5 days) at -20 °C in 50% ethanol until further use.

2.2.4 Immunohistochemistry

2.2.4.1 Immunohistochemistry for FFPE sections

FFPE sections were deparaffinised with xylene (two changes at 5 minutes each), and then rehydrated in graded ethanol concentrations (100% ethanol for 10 minutes followed by 10 minutes in 70% ethanol). Endogenous peroxidase activity was quenched with 0.5% hydrogen peroxide (H_2O_2) in methanol for 10 minutes, and sections washed with TBS (3 times) for 2 minutes each. Antigen retrieval was then conducted by microwaving the sections in 0.01M citrate buffer at pH 6 for 25 minutes and followed by cooling gently under tap water for 3 minutes and washed with TBS (3 times) for 2 minutes each. Next,

avidin was added over the sections on the slides and left for 20 minutes, followed by 3 washes with TBS for 2 minutes each, then biotin was added over the sections on the slides and left for 20 minutes and washed 3 times in TBS for 2 minutes each. Sections were then covered with blocking media (DMEM + 1% BSA + 20% FBS) for 1 hour, which was subsequently drained off and the primary antibody added to the sections on the slides for overnight incubation at 4°C (details of primary antibodies are shown in table 2.2). After 3 washes with TBS for 5 minutes each, biotinylated secondary antibody was applied to the sections for 30 minutes at room temperature (details of secondary antibodies are summarised in table 2.2) followed by 3 TBS washes, 5 minutes each. Avidin-Biotin Complex (ABC) was prepared according to the manufacturer's protocol (1 volume ABC to 75 volumes of TBS) and applied to the sections for 30 minutes, and after 3 TBS washes, 5 minutes each, staining was visualized by incubation of sections with 3, 3′ – diaminobenzidine (DAB chromogen) in DAB substrate buffer for 5 minutes followed by 1 minute TBS wash and 5 minutes wash under running tap water. For counterstaining, sections were immersed in Mayer's haematoxylin for 1 minute, and then excess Mayer's haematoxylin was removed by washing under running tap water for 5 minutes. Sections were then, in the fume hood, dehydrated in graded ethanol concentrations (70% ethanol for 10 minutes followed by 10 minutes in 100% ethanol) and immersed in xylene (two changes at 5 minutes each) before mounting with Di-n-butylPhalate in Xylene (DPX) with a coverslip and left to dry prior to visualising stained sections under a light microscope.

2.2.4.2 Immunostaining for epidermal sheet using anti-p53 anti-bodies

Epidermal sheets which had been fixed in 10% NBF were microwaved for antigen retrieval in 0.01 M citrate buffer at pH 6 for 10 minutes. Then samples were transferred immediately to 12-well plate containing TBS at 4 ˚C and placed on a shaker (Wesbart IS89 multi shelf shaker incubator) at 150 rpm and washed 3 times (2 minutes each). Endogenous peroxidase activity was quenched with 0.5% H_2O_2 in methanol for 20 minutes on a shaker at 150 rpm followed by 3 washes with TBS for 2 minutes each. To increasing the tissue permeability, samples were incubated with 0.05% saponin on a shaker at 150 rpm for 30 minutes and then washed 3 times for 2 minutes each in TBS. Subsequently, samples were immersed in blocking medium (DMEM + 3% BSA + 15% FBS) on a shaker at 150 rpm for 1 hour at room temperature, and after draining off the blocking medium, overnight incubation at 4 °C with anti-p53 antibody (DO-7) (DAKO, UK)

at 1:100 dilutions in blocking solution was conducted. On the next day, the tissue was washed with TBS 3 times (5 minutes each) followed by one hour incubation with biotinylated polyclonal rabbit anti-mouse secondary antibody (DAKO, UK) (1:200 in blocking medium) on a shaker at 150 rpm at room temperature before the sections were washed 3 times with TBS for 5 minutes each. ABC complex was prepared according to the manufacturer's protocol (1 volume ABC to 75 volumes of TBS) and the samples incubated with ABC on a shaker at 150 rpm for 45 minutes before washing 3 times for 5 minutes each in TBS. Samples were then incubated with DAB chromogen in DAB substrate for 5 minutes on a shaker at 150 rpm and washed in TBS for 1 minute followed by 3 separate washes in water 2 minutes each. Finally, samples were mounted with Vectashield onto microscope slides for visualising the immunostained epidermis under a light microscope.

Table 2.2 Antibodies used for immunostaining and the final concentration of each antibody that was applied to the tissue sections. Rows 1-5 are primary antibodies while rows 6 & 7 are secondary antibodies.

	Antibody	Isotype/s	Concentration
1	Monoclonal mouse anti- p53 (DO-7)	IgG2b	1:100
2	Monoclonal mouse anti- beta-catenin (E-5)	IgG1	1:100
3	Monoclonal rabbit anti-CD4	IgG	1:100
4	Monoclonal mouse anti-CD8	IgG1	1:20
5	Monoclonal mouse ani-FOXP3	IgG2a	1:50
6	Polyclonal rabbit anti-mouse immunoglobulin biotinylated	IgG	1:200
7	Polyclonal swine anti-rabbit immunoglobulin biotinylated	IgG	1:200

2.2.5 Haematoxylin and Eosin (H&E) staining

FFPE tissue sections were dewaxed in xylene (two changes at 5 minutes each), then rehydrated through graded ethanol (100% ethanol for 10 minutes followed by 10 minutes in 70% ethanol). After that, slides were immersed in Mayer's haematoxylin for 5 minutes and subsequently washed under running tap water for 5 minutes. Slides were then immersed in eosin for 5 minutes, followed by a brief rinse under tap water before dehydrating the sections in graded ethanol (70% ethanol for 10 minutes followed by 100% ethanol for 10 minutes) and xylene (twice for 5 minutes each) and finally mounting in DPX and the coverslip were applied.

2.2.6 Section imaging

H&E stained and immunostained vertical tissue sections were visualised and imaged using an Axiocam camera connected to a Nikon Eclipse 80i light microscope. Images were taken at different fold magnifications (2, 4, 10 and 40x). As 40x images were used for

quantification of various features (e.g. dysplasia, immunopositivity, etc.), 5 different images at 40x were taken randomly over different areas of the tissue section on each slide to ensure appropriate representation of the tissue section. Epidermal sheets with positive p53 staining that represented p53 immunopositive patches (PIPs) were also imaged at 4x to measure the size of the PIP in mm^2 and were visualised at 40x using Nikon Eclipse 80i light microscope and S10 stage counting grid, to count the cells within the PIPs.

2.2.7 Image analysis

Representative images from H&E and immunostained tissue sections and epidermal sheets were analysed using ImageJ software.

2.2.7.1 H&E stained images

ImageJ software was used for grading and measuring the area of dysplasia within H&E stained lesions. To ensure consistency of measurements, the first step in the analysis was to set a scale bar using a 1 mm image slide and all the downstream measurement were referred to this scale bar. Images of 2x magnification were used to measure the dysplastic area (mm^2) and 40x images were used to count dysplastic cells within the lesions. In the analysis, an area equal to 0.056 mm^2 (which represents the image size when opened with ImageJ software) was selected as the unit area for 40x images. For cell counting, images were firstly converted into binary images of black and white and the ImageJ threshold function was used for brightness adjustment to ensure the best visualisation of individual cells for cell counting. To estimate the total number of dysplastic cell within the lesion, the mean of the dysplastic cell count within the 0.056 mm^2 for each of 5 H&E stained images at 40x magnification was multiplied by the total lesional area (measured in 2x images). Training in the recognition of dysplastic keratinocytes was conducted prior to counting, with agreement between Prof Eugene Healy (Professor of Dermatology), Dr Chester Lai (Dermatology registrar, Southampton General Hospital), Dr Jeffery Theaker (Consultant Histopathologist, Southampton General Hospital) and the author (myself) on the criteria to define dysplastic cells; the latter relied on the combined experience of the above individuals and literature reviews (Boyd et al., 2001, Castano et al., 2002, Yanofsky et al., 2011) and the author was only permitted to undertake the analysis of dysplastic cells when it was agreed that the author was sufficiently experienced in the recognition and grading of dysplasia. Note that in the process of protocol optimisation, ImageJ

assisted manual counting was also conducted, with the aim to compare the results of manual and the automated cell counting. For ImageJ based manual counting, ImageJ; plugins > analysis > cell counter services was used because this enabled counting targeted cells only once through marking and counting cells simultaneously, thus, eliminating the possibility of cell re-counting.

2.2.7.2 p53, CD4, CD8 and FOXP3 immunostained images

Using the same methodology as for counting dysplastic cells, ImageJ software was used to count p53, CD4, CD8 and FOXP3 positively stained cells within the lesions. The same number of images used for quantification of dysplastic features as described above (i.e. one 2x image and 5 images at 40x per slide) were used. Results were expressed as percentage of immunopositive lesional cells; in the case of p53 immunostaining, this was calculated as the mean number of immunopositive cells for a given staining over the mean number of dysplastic cells within the lesion, whereas for CD4, CD8 and FOXP3 staining it was calculated as the mean number of immunopositive cells over the mean number of immune cells within the immune infiltrate. Note that p53 positive cells within PIP were counted manually under light microscope (Nikon Eclipse 80i) using S10 stage counting grid, and 40x lenses, however, ImageJ software was used to measure the area of the PIPs in mm^2 using 4x images.

2.2.7.3 beta catenin stained images

Due to the complexity and the heterogeneity of beta catenin staining, it was difficult to assess staining quality using ImageJ. Consequently, using the principles adapted from Brasanac et al. (2005), manual assessment was conducted to measure the intensity of both normal and abnormal staining. Both staining intensity and cellular localisation of the beta catenin protein were estimated. Comparison of the auther's results with those of two other observers (Prof Eugene Healy and Dr Chester Lai) who also examined the quality of the staining separately on 5 AK sections indicated a high level of agreement between all three observers.

2.2.8 Tissue microdissection

All equipment was cleaned with 70% ethanol and nucleaseliminator unless stated otherwise. All tips were filtered. All microcentrifuge tubes were autoclaved. All disposable items were only used once.

2.2.8.1 Laser capture microdissection (LCM)

Laser capture microdissection of vertical tissue sections was conducted to isolate lesional dysplastic and/or cancerous tissue for DNA analysis. Sections on LCM slides were stained with cresyl violet acetate as follows. Sections were deparaffinised in xylene for 1 minute and transferred to 75% ethanol for 30 seconds. This was followed by immersing the slides in 1% cresyl violet acetate solution for 1 minute before the tissue/slide was washed with 100% ethanol for 1 minute. Stained slides were incubated at 37 °C until they were dry (i.e. left for several hours). To reduce the risk of tissue contamination, the incubator was cleaned with nucleaseliminator and as possible, only the targeted samples were kept in the incubator with no other samples.

Dried, cresyl violet acetate stained slides were then microdissected using an AS LMD laser microdissection microscope (Leica Microsystems). The slides were placed inverted within the microscope, thus, the tissue sections faced downward. Areas of interest, including dysplastic and/or cancerous tissue and adjacent non-lesional normal looking skin were demarcated (using the laser microdissection microscope software) on each tissue section and dissected away from the remaining tissue using the laser beam. The dissected tissue was collected separately into sterile microcentrifuge tubes containing 180 µl of the ALT lysis buffer from a QIAamp DNA FFPE Tissue Kit (Qiagen). Lesional areas were collected separately from the adjacent non-lesional normal looking skin. To decrease the risk of DNA degradation, the maximum laser dissection time did not exceed 45 minutes/ slide (Xu et al., 2008).

2.2.8.2 PIP microdissection

PIPs were identified in wholemount epidermis (which had been stained for p53 protein) according to Ren et al. (1997) and De Graaf et al. (2008) as uninterrupted clusters of ≥ 10 keratinocytes with homogenous and strong nuclear staining for p53 protein. Disposable sterile straight dissecting needles (Thomas Scientific) were used for microdissection of the

PIPs under a GXM-XTL3101, Stereo Zoom Microscope (GT vision), while the edge of the epidermis was held with a fine pointed forceps. A new straight dissecting needle (Thomas Scientific) was used for each PIP and the fine pointed forceps was washed with 70% ethanol and then with nucleaseliminator in between each sample. Dissected PIPs and the surrounding p53 negative epidermis were collected into separate microcentrifuge tubes each containing 180 µl of ALT lysis buffer from QIAamp DNA FFPE Tissue Kit (Qiagen) prior to DNA extraction from the samples. In general, the DNA extraction (as per 2.2.9 below) was performed immediately after the dissection of the target tissue.

2.2.9 DNA extraction

DNA was extracted from microdissected dysplastic and/or cancerous areas, p53 immunopositive patches (PIPs) and adjacent (non-lesional and/or non-PIP) normal looking skin samples using a QIAamp DNA FFPE Tissue Kit according to the manufacturer's protocol as follows. The tissue was lysed by incubating the microcentrifuge tubes from the previous step (2.2.8) containing 180 µl of ALT lysis buffer from the QIAamp DNA FFPE Tissue Kit and the dissected tissue sections (the amount of dissected tissue per sample did not exceed 25 mg as recommended by the QIAamp kit) with 20 µl of proteinase K (stock concentration 20 µg/µl) in a water bath at 45 °C. The tissue was assessed for complete lysis every 4 to 5 hours and whenever required, additional doses of 20 µl of proteinase K was added to the mix (up to a total of 80 µl of proteinase K was added for some samples), and each addition of proteinase K was followed by brief vortexing of the microcentrifuge tubes and centrifuging at 600 RCF for several seconds. Once the tissue was digested, 200 µl of AL buffer (provided from the QIAamp DNA FFPE Tissue Kit) and 100% ethanol (200 µl) was added to the mix in the microcentrifuge tubes, then briefly vortexed to homogenise the mix and centrifuged at 600 RCF for several seconds. Subsequently, the mixture was transferred to QIAamp tubes containing a filter (provided with the QIAamp DNA FFPE Tissue Kit). These tubes with the mixture added were centrifuged at 6,000 RCF for 1 minute thus trapping the DNA on the filter and allowing impurities (chemicals and proteins) to pass through the filter and be discarded. To ensure adequate removal of the impurities and optimise the purity of the DNA, the filtered tubes were washed twice as follows. The first wash was performed by the addition of 500 µl of AW1 to the filter tube which was then centrifuged at 6,000 RCF for 1 minute and followed by the second wash with the addition of 500 µl of AW2 to the filter tube which was then centrifuged for 3

minutes at 20,000 RCF. To ensure the complete removal of ethanol and washing buffers that might inhibit any downstream experiments, these filter tubes had a final centrifugation step conducted at 20,000 RCF for 3 minutes. DNA which was attached to the filter within these tubes was then eluted by the addition of AE elution buffer (provided within the kit) to the filter and left for 5 minutes at room temperature before the tubes were centrifuged for 1 minute at 6,000 RCF. The amount of the added AE buffer varied according to the required amount and concentration of DNA for the downstream analysis, thus, between 25 - 65 µl of AE buffer was used for the target enriched sequencing and whole exome sequencing analysis respectively. The DNA samples were then stored at -20 °C until further use.

2.2.10 DNA quantity assessment

2.2.10.1 Nano-drop absorbance spectrophotometer

Although Nanodrop D-1000 (Nanodrop Technologies, US) absorbance spectrophotometer was used mainly for assessing the quality of DNA, this also provided a preliminary estimation of the DNA concentration. As per the manufacturer's protocol, 2 µl of autoclaved deionised water was loaded firstly into the pedestal of the machine to initiate the measurement (as a reference). Subsequently, 2 µl of the elution buffer that used to elute DNA samples (AE buffer) was loaded to familiarize the programme with the buffer. Then 2 µl of the test DNA samples were loaded consequently into the spectrophotometer. After each loading (including loading of deionised water, buffer and DNA sample), the pedestal was wiped with clean dry tissue paper to prevent inter-sample contamination. DNA concentration was measured in ng/µl and the A260/A280 and A260/A230 nm ratios were also recorded to assess DNA quality. The normal range that indicates pure DNA is an A260/A280 ratio between 1.8 - 2 and readings beyond this range indicate protein contamination, while the normal range of A260/A230 ratio is 2 - 2.2 and results out of this range indicate contamination with solvents, salts and inorganic compounds (Ludyga et al., 2012).

The basis of the Nanodrop technique depends on the ability of DNA to absorb ultraviolet radiation (UV) at specific wavelengths. Consequently, the amount of UV that is absorbed by the DNA within the samples can be recorded and used for the process of measurement of DNA amount. Although Nanodrop is widely used to quantify DNA within samples, its

specificity can be an issue. This is because the presence of contamination by proteins and other chemicals which absorb light at the same wavelengths may result in over-estimation of the amount of DNA in the tested samples.

2.2.10.2 Qubit 2.0 Fluorometer

As accurate DNA concentration is crucial for next generation sequencing, Qubit 2.0 fluorometer (which specifically recognises DNA within the tested samples) was used. Qubit 2.0 fluorometer relies on the hybridisation of fluorescent dye with the targeted molecule for the purposes of quantification. Different dyes can be used that bind specifically to the targeted molecule (DNA, RNA or protein) and the fluorescence only becomes detectable upon this specific binding. Equilibrium of binding of the fluorescence dye to double stranded DNA is reached in less than 2 minutes and the amount of produced fluorescence is directly proportional to the DNA concentration within the sample. Thus, the resultant fluorescence can be converted into a numerical result that provides the DNA concentration in ng/µl.

Qubit 2.0 fluorometer was performed according to the manufacturer's instructions. For DNA quantification, the Qubit dsDNA broad range assay on the machine was selected (this can detect DNA concentrations between 100 pg/µl -1 µg/µl). A working solution was prepared by mixing (1 µl x n) of the Qubit dsDNA broad range reagent that contains the fluorescence dye with (199 µl x n) of the Qubit dsDNA broad range buffer [where n is equal to the number of tested DNA samples plus 2 standards (Qubit dsDNA broad range standard 1and 2) provided with the Qubit® dsDNA BR assay kit]. Then, 10 µl from each of the two standards (used as reference) was added separately to two clean autoclaved microcentrifuge tubes each containing 190 µl of the prepared buffer/reagent mix from the previous step. In separate autoclaved microcentrifuge tubes containing 2 µl of each of test DNA samples, 198 µl of the same reagent/buffer mix was added to each of the microcentrifuge tubes containing the tested DNA (2µl) and mixed well. All samples and standards were kept in the dark at room temperature for about 2 minutes before starting the measurements. The microcentrifuge tubes were then loaded into the machine in consecutive manner, starting with the standards 1, 2 microcentrifuge tubes and followed by the test DNA samples. The DNA concentration was expressed as ng/µl.

2.2.11 DNA quality assessment

In addition to Nano-drop Absorbance Spectrophotometer (section 2.2.10.1) giving an indication of the DNA purity, a multiplex PCR was performed to determine whether amplification of DNA up to 600bp sizes could be obtained from the DNA samples.

2.2.11.1 Polymerase chain reaction (PCR)

The master mix for PCR was prepared on ice. Two PCRs were performed to assess DNA quality, one which amplified the *MC1R* gene and the other was a panel (multiplex) PCR which amplified specific DNA fragments of lengths 100bp, 200bp, 300bp, 400bp and 600bp. The reaction conditions for the *MC1R* gene PCR were adapted from previous work in the department (Robinson and Healy, 2002). The reaction conditions for the multiplex PCR was adapted from in-house set of guidelines (cancer genomic guidelines, biomedical PCR DNA quality assessment, University of Southampton, UK). Moreover, a third PCR on the *KNSTRN* gene was also performed in order to undertake Sanger sequencing of this gene; the *KNSTRN* gene primer pairs were designed using (NCBI/BLAST) (www.ncbi.nlm.nih.gov/tools/primer-blast) tool. PCR was conducted using a GeneAmp®, PCR system 9700 thermocycler from Applied Biosystems. The reaction conditions for each of the three PCRs are detailed in table 2.3, and the primer sequences and PCR thermal conditions in tables 2.4 and 2.5 respectively. The PCR products were kept at 4°C until further analysis using agarose gel electrophoresis.

Table 2.3 Master mix preparations for *MC1R* gene PCR, *KNSTRN* gene PCR and panel (multiplex) PCR.

Master mix	**PCR**		
Reagents	***MC1R* gene PCR (final concentration in parentheses)**	***KNSTRN* gene PCR (final concentration in parentheses)**	**Panel (Multiplex) PCR (final concentration in parentheses)**
MyTaq reaction buffer	5 µl of 5x buffer (1x buffer)	5 µl of 5x buffer (1x buffer)	5 µl of 5x buffer (1x buffer)
Forward and reverse primers	0.6 µl of 50 µM concentration (0.48 µM each primer)	0.6 µl of 50 µM concentration (0.48 µM each primer)	0.125 µl of 50 µM concentration (0.25 µM each primer)
DNA template	2 µl of 15 ng/µl (30ng)	2 µl of 15 ng/µl (30ng)	2 µl of 15 ng/µl (30ng)
MyTaq Red DNA Polymerase	1 µl of 5 unit/µL (5 unit)	1 µl of 5 unit/µL (5 unit)	1 µl of 5 unit/µL (5 unit)
Nuclease free water	15.8 µl	15.8 µl	15.75 µl
Final volume	25µl	25µl	25µl

Table 2.4 PCR primer pairs.

primer ID	Primer pair	Primer sequence	Length in base pairs	Amplified fragment length in base pairs	Annealing temperature (Tm)	CG content %
TBXAS1/X9U	Forward	GCCCGACATTCTGCAAGTCC	20	100	63.75	60
TBXAS1/X9L	Reverse	GGTGTTGCCGGGAAGGGTT	19	100	65.06	63
RAG1/X2U	Forward	TGTTGACTCGATCCACCCCA	20	200	63.28	55
RAG1/X2L	Reverse	GAGCTGCAAGTTTGGCTGAA	20	200	60.10	50
PLZF/X1U	Forward	TGCGATGTGGTCATCATGGTG	21	300	64.72	52
PLZF/X1L	Reverse	GTGTCATTGTCGTCTGAGGC	20	300	57.63	55
AF4/X11U	Forward	CCGCAGCAAGCAACGAACC	19	400	65.90	63
AF4X11L	Reverse	GCTTTCCTCTGGCGGCTCC	19	400	65.31	68
AF4/X3U	Forward	GGAGCAGCATTCCATCCAGC	20	600	63.87	60
AF3/X3L	Reverse	CATCCATGGGCCGGACATAA	20	600	64.51	55
***MC1R* gene**	Forward	ATGGCTGTGCAGGGATCCCC	20	954	59	54
***MC1R* gene**	Reverse	TCACCAGGAGCATGTCAGCA	20	954	62	50
***KNSTRN* gene**	Forward	CCTTTCGCTAGGTCTGGCTC	20	317	60.1	60
***KNSTRN* gene**	Reverse	CATTCCGTCCTTCCTCCCAT	20	317	59.1	55

Table 2.5 PCR reaction conditions

		***MC1R* gene PCR**		***KNSTRN* gene PCR**		**Panel (Multiplex) PCR**	
		Time	Temperature	Time	Temperature	Time	Temperature
Denaturation		5 minutes	95 °C	5 minutes	94 °C	5 minutes	95 °C
35x Cycles	**Denaturation**	1 minutes	94 °C	30 seconds	94 °C	30 seconds	95 °C
	Annealing	1 minutes	62 °C	30 seconds	60.4 °C	30 seconds	60 °C
	Extension	2 minutes	72 °C	30 seconds	72 °C	30 seconds	72 °C
Final extension		7 minutes	72 °C	7 minutes	72 °C	15 minutes	72 °C

2.2.11.2 Agarose gel electrophoresis of genomic DNA and PCR product

Agarose gel electrophoresis was used to assess the quality of the genomic DNA (where a high band on the gel indicates limited degradation of the DNA) and also to identify the size of PCR products. A 2% w/v agarose gel was prepared by dissolving 2 grams of agarose crystal (Sigma) in 100 ml of 1x TBE buffer (Fisher Scientific) by heating in a microwave. Then, the agarose gel was left to cool at room temperature until its temperature reached ~40°C before 3 µl of Nancy 320 dye (Sigma) was added to 100 ml of the agarose gel to assist in genomic DNA/PCR product visualisation. Immediately after addition of the dye, the gel was poured into a mini-gel former and left to solidify after the insertion of the appropriate comb (to generate wells for loading of samples). Once the gel had solidified, the comb was removed and the gel was immersed gently and completely in a gel tank

containing 1x TBE buffer. DNA samples were mixed with DNA loading dye (Bioline, UK), in a ratio of 5x of DNA: 1x loading dye (v/v), and then added to the gel wells. For PCR products, the loading dye was already included within the 5x MyTaq reaction buffer reaction buffer supplied with MyTaq™ Red DNA Polymerase kit, therefore, appropriate volumes of PCR products (according to the size of the gel wells) were loaded to the gel. A DNA ladder (HyperLadder™ 100bp) from Bioline was run simultaneously within the same gel for identification of the size of the DNA fragments or PCR products on the gel. The gel was then electrophoresed at 40V constant voltage for 1 to 2 hours depending on the expected sizes of the DNA fragments. For visualisation of the electrophoresed DNA / PCR products on the gel, the DigiDoc-It® Imaging System was used.

2.2.12 Next generation sequencing (NGS)

Whole exome sequencing (WES) and target enriched sequencing were performed in the Wellcome Trust Centre for Human Genetics, Oxford laboratories, Oxford, UK. For WES, the SureSelect V5 (Agilent V5) exome capture kit was used for library preparation, while the Illumina HiSeq2000 system was used for sequencing. The TruSeq custom amplicon kit v1.5 (Illumina) kit was used for target enriched sequencing library preparation, while the Illumina MiSeq system was used for sequencing. An overview of WES and target enriched sequencing is illustrated in figure 2.1 and is provided in the subsequent paragraphs for information.

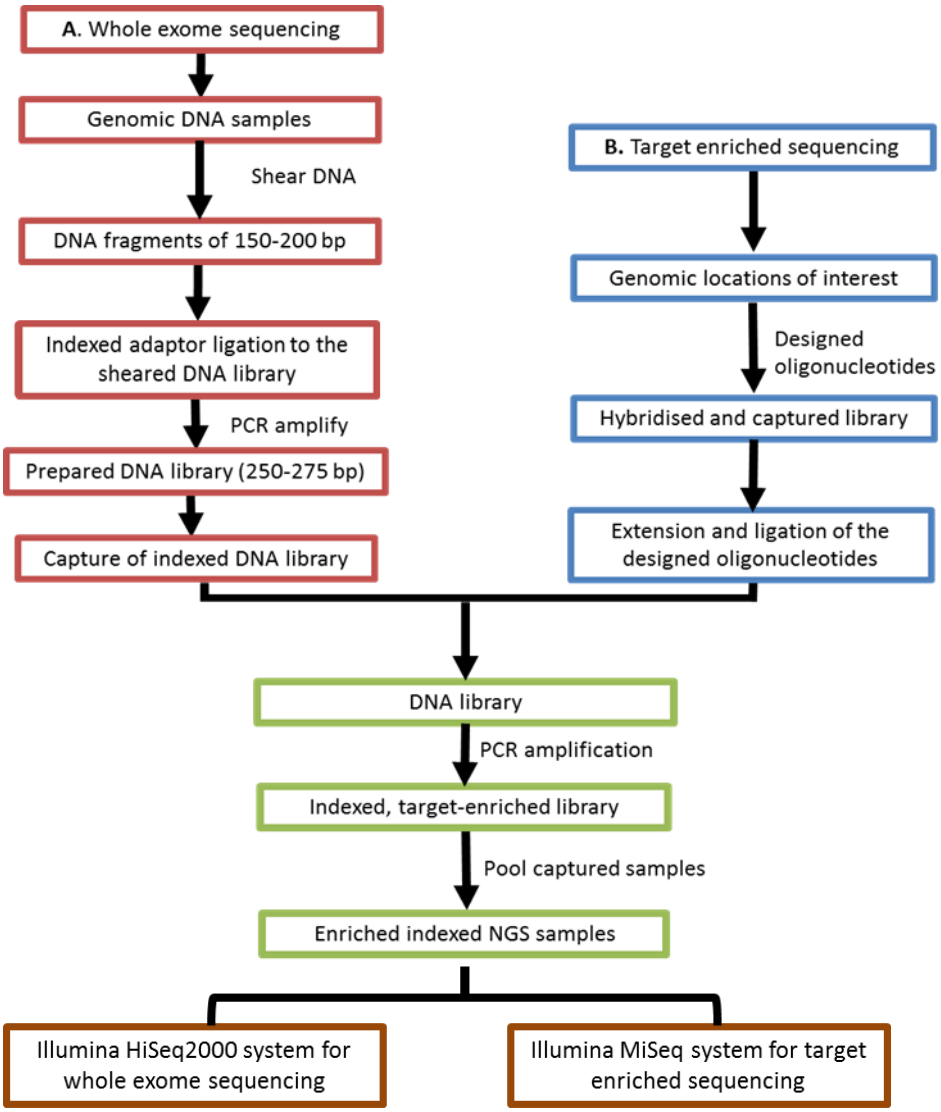


Figure 2.1 Flowchart on the steps of next generation sequencing. **A:** Whole exome sequencing workflow. **B:** Targeted enriched sequencing workflow. NGS = Next generation sequencing. bp = base pair.

2.2.12.1 Overview of whole exome sequencing (WES)

Using the SureSelect V5 library preparation kit from Agilent, (Agilent Technologies, 2014) (http://www.agilent.com/cs/library/usermanuals/Public/G7530-90000.pdf), genomic DNA library is prepared by shearing the DNA samples to the size of 150 - 200 bps using a Covaris focussed ultrasonicator (ultrasound based DNA shearing) and the quality of the sheared library is assessed using a 2100 Bioanalyser DNA 1000 assay. This is followed by a DNA ends repair step (in the presence of dNTPs, DNA polymerase and DNA polymerase kinase) which results in the production of blunt-ended DNA for efficient ligation of the library to the adaptors. The sheared/repaired DNA library is then purified using AMPure

XP beads (provided with the kit) in a magnetic separator stand. Subsequently, the 3' prime end of the DNA fragments are adenylated through the addition of a mono 'Adenosine' nucleotide to the 3' ends of the blunt DNA fragments to allow ligation to the indexing adapters that are designed to have a projection of single 'Thymine' nucleotide that binds the 'Adenosine' nucleotide on DNA fragments during the adapter ligation reaction. The indexed pre-capture library is then amplified and the quality of the amplified library assessed using the 2100 Bioanalyser DNA 1000 assay. The ideal DNA fragment size should be around 250 to 275 bps at this stage. The next step is capturing the indexed genomic DNA library by hybridisation of the pooled, indexed DNA library with SureSelect capture library oligonucleotides that are designed to capture all the exons in the genomic DNA.

The capture library baits provided within the kit are biotinylated, thus the DNA target library hybridised to the biotinylated capture library can be purified using streptavidin magnetic beads in a magnetic separator strand, and the unbound oligonucleotides washed away. A second amplification step to the indexed, pooled, captured DNA library is performed and the resulting products then purified using the AMPure XP beads and assessed for their quality using 2100 Bioanalyzer high sensitivity DNA assay.

The next step is cluster generation, which involves the use of the TruSeq PE cluster kit v3 to generate millions of copies from the indexed library (10 million single strand DNA molecules per cm^2) for sequencing using the HiSeq2000 system. Cluster generation relies on the complementary binding of the adapters on single stranded DNA from the library to primers on the surface of a flow cell to allow solid phase bridge amplification of the DNA (in the presence of dNTPs and polymerase enzyme). The resultant double stranded molecule is then denatured and the process of bridge amplification is repeated to generate millions of clusters of target DNA molecules ready for sequencing.

Illumina HiSeq2000, v3, 100 bp, paired end platform (high output run mode; 28 Gb per lane) was used for WES sample sequencing and 5 samples were loaded per lane (the maximum capacity is 8 samples). Illumina sequencing is based on sequencing by synthesis which uses 4 different fluorescently labelled 2'- deoxyribonucleotide triphosphate (dNTP). During each cycle, one dNTP is incorporated onto the sequencing strand that should be complementary to the single DNA strand on the clustered library and the resultant colour

from the fluorescent dye indicates which dNTP is incorporated. This dye is then enzymatically cleaved to allow the incorporation of the next dNTP. Base incorporation relies on the natural competition between the 4 dNTPs (A, T, C and G) which are present in equal amounts for each run (reducing the chance of base misalignment).

2.2.12.2 Overview of target enriched resequencing

Using the TruSeq custom amplicon kit v1.5 (Illumina), (Illumina, 2015) (http://www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/datasheet_truseq_custom_amplicon.pdf). the first step in library preparation involves denaturing of the double stranded DNA, then hybridisation of custom specific oligonucleotides (designed to hybridise to relevant genes) to the target genomic DNA and subsequent removal of unbounded oligonucleotide on a filter plate membrane, thus, the hybridised DNA remains on the membrane and unbounded oligoes are washed away. The hybridised oligonucleotides are then extended by DNA polymerase along the complementary targeted DNA, and then the adjacent ends of the extended oligonucleotides over the targeted DNA are ligated using a DNA ligase. The resultant products of the extension-ligation reaction contain the targeted regions of interest flanked by the oligonucleotides sequences required for PCR amplification. During PCR amplification, the primers are designed to be complimentary to the hybridised oligonucleotides from the previous step and also to be ligated to unique sequences (indexes) that are important for sample identification in multiplex PCR/sequencing reactions. When the PCR is completed, the size of the resultant library should be 350 base pairs for DNA input of 250 base pairs in size. The resultant library is then purified using AMPure XP system (Beckman, UK) and pooled together.

Cluster generation is exactly as for WES (as described above) using flow cell and bridge amplification of the cluster.

For target enriched sequencing, Illumina MiSeq system, v3, read length 300 bps, with output data (15 Gb) was used. As for HiSeq2000 system, the sequencing strategy of MiSeq system relies on sequencing by synthesis (as described for WES above).

2.2.13 Sanger sequencing

The basis of Sanger sequencing relies on selective integration of a chain terminating 2' 3' – dideoxynucleotide triphosphate (ddNTP) to the growing DNA chain using DNA polymerase. The 4 different ddNTPs (ddATPs, ddTTPs, ddCTPs, ddGTPs) included in the reaction are labelled with different radioactive or fluorescent dyes, thus, the incorporated terminating base can be recognised by the automated sequencing machine (capillary electrophoresis). The classical Sanger sequencing method requires a single-stranded DNA template, DNA primer, DNA polymerase, dNTPs for chain elongation, ddNTPs in reaction buffer. Sanger sequencing for PCR products from the *KNSTRN* gene were performed by Source Bioscience, Nottingham, UK using an Applied Biosystems 3730/3730xl sequencer (Applied Biosystems, 2010).

Note; For WES and target enriched sequencing, a mutation was defined as a base pair alteration that was present in the lesional DNA, which would result in an amino acid alteration or deletion or a frameshift that was predicted to have a deleterious effect on protein function according to the following software programs; Phylop, SIFT, Polyphen2, lrt, mutationtaster, gerp++, and Grantham scores.

Mutations were identified using the following approach; the base pair alteration

(I) Should be somatic and non-silent (missense, nonsense, frameshift and splicing) mutations as annotated by Varscan 2.3.3.

(II) Should be annotated as novel or clinical and should not be reported previously as SNP within the 1000 Genomes Project (2012 April release) and dbSNP135 database.

(III) Should have a p value for somatic changes ≤ 0.05 (somatic-p-value is defined as threshold below which read count differences between lesion and normal are deemed significant enough to classify the sample as a somatic mutation or an LOH event).

(IV) Should be present in ≥ 4 of the lesional reads and < 4 of the normal reads.

In addition, in order to avoid false positive identification of mutations, variants that were flagged as or within homopolymers (highly repetitive regions) were excluded.

Chapter 3: Genetic landscape of Actinic Keratoses (AKs)

3.1 Introduction

Studying the genetic alterations within cancers and their potential precancerous lesions is one of the ways that can be used to improve the understanding of the relationship between the two types of lesions (Vogelstein et al., 2013). Various genetic studies have been performed on cSCC and AK looking at base pair changes, microsatellite instability and, in some cases, methylation status.

TP53 mutation and abnormal p53 protein expression are seen in both cSCCs and AKs, with a higher frequency of *TP53* mutation in cSCCs than AKs (Brash et al., 1991, Kramata et al., 2005, Kubo et al., 1994, Nelson et al., 1994, Nindl et al., 2007, Ziegler et al., 1994). In addition to *TP53* mutation within AKs, microsatellite studies of the 17p13 region revealed LOH within the region encoding for *TP53* in 64% of AKs (Rehman et al., 1996). LOH has also been observed on chromosomes 9q, 9p, 17q, 3p and 13q (as well as 17p) in AK (Kushida et al., 1999, Rehman et al., 1994, Rehman et al., 1996). In one study, deletion of the 9p21 region (encompassing *CDKN2A*) was only detected within cSCC, but not in AK and the authors suggested an importance for inactivation of *CDKN2A* in the process of malignant transformation (Mortier et al., 2002). However, in a later study, *CDKN2A* (encoding p16INK4a / p14ARF) sequencing showed missense and nonsense mutations in *CDKN2A* in 12% of AKs (Nindl et al., 2007) consistent with inactivation of *CDKN2A* within AKs. Moreover, mutational analysis of *CDKN2A* in 26 AK samples in further study revealed a missense mutation at codons 65, 71 and 184 of *CDKN2A* (Kanellou et al., 2008).

The mutational status of common oncogenes has also been investigated in AKs. The *RAS* family (*HRAS, KRAS* and *NRAS*) genes are known oncogenes that are frequently seen mutated in human cancer (Fernandez-Medarde and Santos, 2011). Analysis of *RAS* mutation status within AKs revealed *RAS* mutations at codon 12 of *KRAS* (Nindl et al., 2007, Spencer et al., 1995, Zaravinos et al., 2010) and codons 12, 13, and 61 of *HRAS* (Spencer et al., 1995). Another oncogene is c-*MYC* gene, which plays a role in cellular proliferation and is seen to be altered within a range of human cancers (Gostissa et al., 2009) including cSCC (amplification of *c-MYC* was detected in 50% of cSCCs from renal transplant patients) (Boukamp, 2005). *c-MYC* has also been investigated in AKs. Using fluorescence in situ hybridization (FISH), multiple copies of c-*MYC*, indicating gain and amplification of the gene have been detected in 35% and 63% of AKs and cSCCs

respectively (Toll et al., 2009). Copy number changes of the epidermal growth factor receptor (*EGFR*) gene within AKs have also been investigated. *EGFR* gene encodes for a member of the C-erbB receptors of the transmembrane type I receptor tyrosine kinase family and *EGFR* activation is associated with epithelial growth, cell proliferation and tumour progression (Pastore et al., 2008). According to Toll et al. (2010), 52% of AK and 77.1% of a cSCC have multiple copies of the *EGFR* gene. The high percentage of genetic alteration in c-*MYC* and *EGFR* within cSCC in comparison to AKs may suggest the involvement of these genes in cSCC pathogenesis (Toll et al., 2009, Toll et al., 2010). However the fact that more than 50% of AKs had these alterations could equally support the concept that alterations in these genes are relevant to development of AKs. The *E-cadherin* gene belongs to the *cadherin* gene family and plays a major role in cell–cell interactions in epithelium (Takeichi, 1995). Reduced expression of E-cadherin protein has been associated with metastatic tendency in cSCC (Koseki et al., 1999). Subsequently, silencing of the *E-cadherin* gene promoter CpG island by hypermethylation was detected in 44% of AKs and 85% of invasive cSCCs (Chiles et al., 2003) which may explain the reduction in the gene expression seen in cSCC and also suggest that this gene may be relevant in AK development.

In addition to the single gene based studies described above, more recently, whole exome sequencing (WES) has been performed to study base pairs alterations within cSCC (Durinck et al., 2011, Li et al., 2015, Pickering et al., 2014, South et al., 2014). The advantage of WES is that it allows one to investigate cancers and precancerous lesions in a more comprehensive way than previous candidate gene studies and prior to commencing the current project in this thesis; WES of AKs has not been published by any other research group.

3.2 Hypothesis and aims

The hypothesis of this study was that actinic keratosis lesions harbour a high mutation burden and harbour mutations of driver genes that are altered in cSCC.

The aim of this chapter was to define the molecular changes within a small group of AKs and to compare the mutation burden and frequency between AKs and cSCCs from previous studies using WES.

3.3 Materials and Methods

3.3.1 Tissue sample

Archived FFPE AK blocks were collected from the Histopathology department at Southampton General Hospital. Fresh surgically resected redundant normal breast skin was obtained from patients undergoing mastectomy in the Surgical department at the Princess Anne Hospital, Southampton which then underwent formalin fixation and paraffin embedding immediately in the Histochemistry Research Unit (HRU) in Southampton General Hospital. Blood DNA samples were obtained from patients as part of an investigation on skin cancer being conducted by Dr Chester Lai in Dermatopharmacology, University of Southampton. The study was conducted under local research ethics committee approval and a signed consent form was obtained from all subjects for the use of the tissue.

3.3.2 Haematoxylin & Eosin (H&E) staining

5 µm tissue sections of AK and normal breast skin were stained for H&E as described in section 2.2.5. H&E slides of AKs were used to identify the dysplastic area within the AKs and the number of dysplastic cells, and to grade dysplasia within the lesion.

3.3.3 Immunohistochemistry

AK sections (5 µm thick each) were stained using the different antibodies as described in section 2.2.4. AK sections were immunostaining using antibodies that target different cellular proteins (anti-p53, anti-beta catenin, anti CD4, anti CD8 and anti FOXP3 antibodies (as detailed in table 2.2). The aim of immunostaining was to help delineate the dysplastic area (e.g. looking at p53 and beta catenin stains alongside H&E stain) and to assist in recognition of non-dysplastic cells (predominantly immune cells which were stained by CD4, CD8 and FOXP3), thus aiding microdissection of the dysplastic keratinocytes from samples selected for genetic investigations.

Images at different fold magnifications (2x, 10x and 40x) were taken for each stained AK section using an Axiocam camera connected to a Nikon Eclipse 80i light microscope. For each slide, 1 image at 2 and 10 fold magnification and 5 images representing 5 different areas within the section at 40 fold magnification were included.

3.3.4 Laser capture microdissection and cresyl violet acetate staining

To enrich for dysplastic lesional keratinocytes within AK samples, 10 µm thick tissue sections from 5 selected FFPE AK blocks were laser capture microdissected after being stained with cresyl violet acetate as described in section 2.2.8.1. The matched adjacent non-dysplastic perilesional normal looking skin tissue was also laser capture microdissected and used as control. According to the sample size, as estimated from the reference H&E slides, dysplastic keratinocytes were isolated from between 40 – 55 AK sections and the dissected tissue collected in microcentrifuge tubes containing 180 µL of ALT lysis buffer for subsequent DNA extraction.

3.3.5 DNA extraction

Genomic DNA was extracted from the lesional (dysplastic) and matched adjacent normal looking skin tissue of AKs, and from breast skin samples and blood samples.

As outlined in section 2.2.9, genomic DNA was purified using the QIAamp DNA FFPE Tissue Kit (Qiagen) according to the manufacture's recommendation. Target DNA concentration for the WES analysis was 30 - 50 ng/µl (total of ≥ 500 ng).

3.3.6 Assessment of DNA concentration within the samples

A Qubit 2.0 fluorometer (Invitrogen) was used for quantification of genomic DNA, as described in section 2.2.10.2, on DNA isolated from the AK, breast and blood samples.

3.3.7 Assessment of DNA quality

Using Nanodrop D-1000 (Nanodrop Technologies) absorbance spectrophotometer, the optical density in the form of A260/A280 and A260/A230 ratios was measured for each DNA sample as detailed in section 2.2.10.1; this measurement is important to assess if there was any contamination of DNA samples with organic or non-organic substances. DNA fragmentation was assessed (as detailed in section 2.2.11) using agarose gel electrophoresis and PCR. PCR for the 1Kb coding region of the *MC1R* gene and a panel (multiplex) PCR using different primer pairs that amplify different DNA fragment lengths (100, 200, 300, 400 and 600 bps) were used. The primers and reaction conditions for *MC1R* gene and panel PCRs were as detailed in the section 2.2.11.1. The resultant bands were visualised on agarose gel electrophoresis as detailed in section 2.2.11.2.

3.3.8 Whole exome sequencing (WES)

Genomic DNA (N = 10, i.e. 5 from lesional DNA and 5 from the matched adjacent normal looking skin) from 5 different AKs was selected for WES. In brief, between 825 – 5820 ng of DNA per sample was sent on dry ice to the Wellcome Trust Centre for Human Genetics, University of Oxford for WES. Agilent SureSelect V5 Human All Exon 50 Mb (Agilent) was used for target capture and library preparation and Illumina HiSeq 2000 platform for DNA sequencing according to the manufacturers' instructions as described in section 2.2.12.1.

3.3.9 Read mapping and variant calling

Whole exome paired-end sequencing data were aligned against the human GRCh37/hg19 (Genome Reference Consortium 37) using the Novoalign software (novoalignMPI V2.08.02, Novocraft Technologies, Selangor, Malaysia). Duplicate reads, resulting from PCR clonality or optical duplicates, and reads mapping to multiple locations were excluded from downstream analysis. Depth and breadth of sequence coverage was calculated with custom scripts and the BedTools package (v2.13.2) (Quinlan and Hall, 2010). Germ-line-Tumour paired datasets were analysed to identify single nucleotide variations (SNVs) and small insertion and deletions using Varscan 2.3.3 (Koboldt et al., 2012). The minimum variant allele frequency threshold was set to 20% with a minimum read depth of 4. Variants were filtered using the 'somaticFilter' command to remove clusters of false positives and SNV calls near indels with the same frequency and depth thresholds. Somatic p-values for a Fisher's exact test of read counts for reference and alternative alleles in the two matched samples are provided by Varscan, and a threshold of significance of $p \leq 0.05$ was required for all variants. Variants were annotated with respect to genes and transcripts and filtered using the Annovar software tool (v2012Jun21) (Wang et al., 2010). Finally variants were annotated with functional prediction scores, SIFT (Sorting Intolerant From Tolerant) and Polyphen2 (Polymorphism Phenotyping) Phylop, lrt, mutationtaster, gerp++, and Grantham scores. and non-silent mutations (missense, nonsense and splicing) were retained for analysis. Read mapping and variant calling was undertaken by Dr Reuben Pengelly, Genetic Epidemiology and Genomic Informatics Group, Human Development and Health, Faculty of Medicine, University of Southampton.

3.3.10 Statistics

GraphPad Prism (v6.0) was used to generate figures and to undertake statistical analysis using one way ANOVA and linear regression analysis (statistical support was received from Scott Harris and Ho-Ming (Brian) Yuen, Medical Statistics, Primary Care and Population Sciences, Faculty of Medicine, University of Southampton).

3.4 Results

3.4.1 Patient identification

FFPE AKs were identified with assistance from Dr. Chester Lai as follows; A list of patients whose AK samples were stored in Histopathology, University Hospital Southampton NHS Foundation Trust was generated by searching for 'actinic keratosis and skin' on the Histopathology results database. The electronic results system of University Hospital Southampton NHS Foundation Trust was used to obtain the histopathology reports for each patient and these were assessed before sample selection. A signed consent form was included within the records of all subjects involved in this study. Using the University Hospital Southampton NHS Foundation Trust Histopathology department computerised records system, 69 patients with actinic keratosis were identified. During the process of sample identification, factors such as the size of the lesion were considered; this was mainly to ensure the availability of sufficient tissue for DNA extraction. Patient characteristics and histopathological data are summarised in table 3.1 and appendix table 8.1. Of the 69 samples, 28.9% of the patients were women and 71 % men. The median age of the study population was 79 years with a range between 44 - 93 years. Most of the lesions were on sun exposed skin areas and distributed mostly on the face and scalp.

Table 3.1 Characteristics of patients with AKs used in this study. 69 AK samples from 69 patients were selected from different body sites. 28.9% of patients were females and 71% were males with median age range between 44 – 93 years.

		AK samples	
Gender	**Male**	**Female**	**Total**
	49	**20**	**69**
Site of the lesion			
Scalp	8	1	9
Forehead	10	2	12
Cheek	4	3	7
Ear	8	0	8
Nose	1	2	3
Hand	4	4	8
Forearm /arm	3	5	8
Neck	3	3	6
Chest	4	0	4
Back	2	0	2
Lower leg	2	0	2
Age at biopsy	Range (44-93 years), Median 78 years	Range (54-93 years), Median 81.5 years	Range (44-93 years), Median 79 years

3.4.2 Analysis of actinic keratosis staining

3.4.2.1 Haematoxylin and Eosin (H&E) staining

Examples of H&E staining are shown in figure 3.1 and demonstrated hyperkeratosis, parakeratosis, hyperplasia and dysplasia in the AK samples.

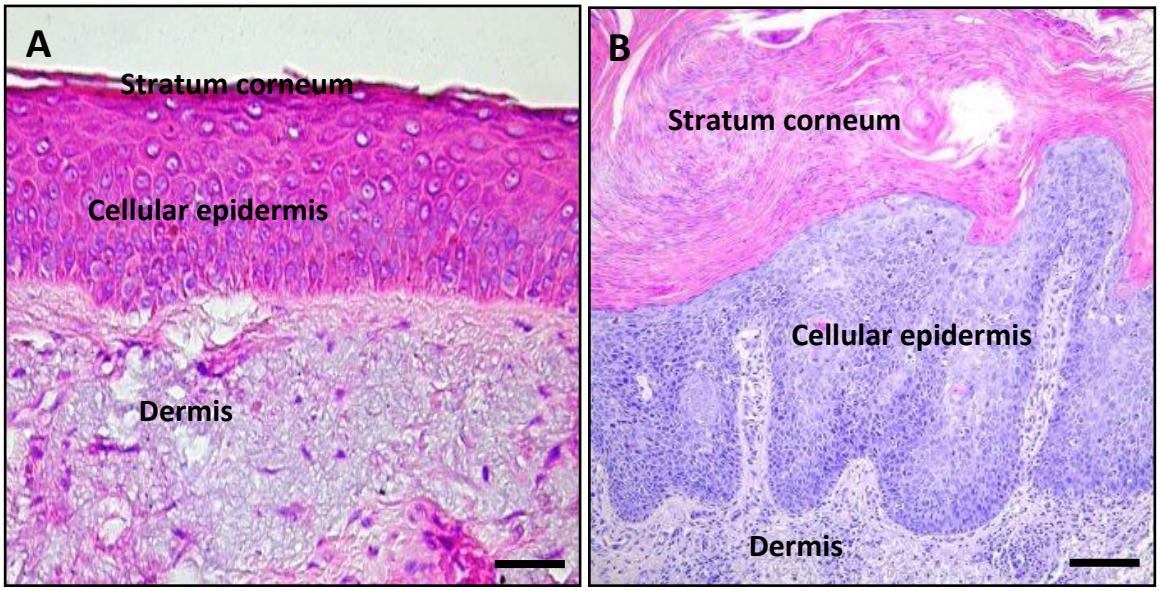


Figure 3.1 H&E staining of actinic keratosis. **A:** Normal skin. **B:** Actinic keratosis section showing hyperkeratosis (increase in the thickness in the stratum corneum layer), parakeratosis (retention of nuclei in the stratum corneum which is normally non-nucleated) and hyperplasia (increase in the number of cells) and dysplasia (cellular abnormality) in the cellular epidermal compartment. 10x represent fold magnification of images. Scale bar= 50 µm.

3.4.2.1.1 Assessment the grade of dysplasia within AK samples

As the main purpose of H&E staining was to grade and quantify the dysplasia within AKs, figure 3.2 shows examples of the parameters that were used to indicate keratinocyte dysplasia. These parameters include; nuclear atypia, loss of polarity, mitotic activity, basal cell layer proliferation, adnexal involvement with dysplastic cells and dyskeratosis. However, AKs also contain other non-dysplastic features such as acantholysis, vacuolization and hyper/parakeratosis which are not relevant in AK diagnosis. The dysplastic/non dysplastic features of AK are illustrated in figure 3.2

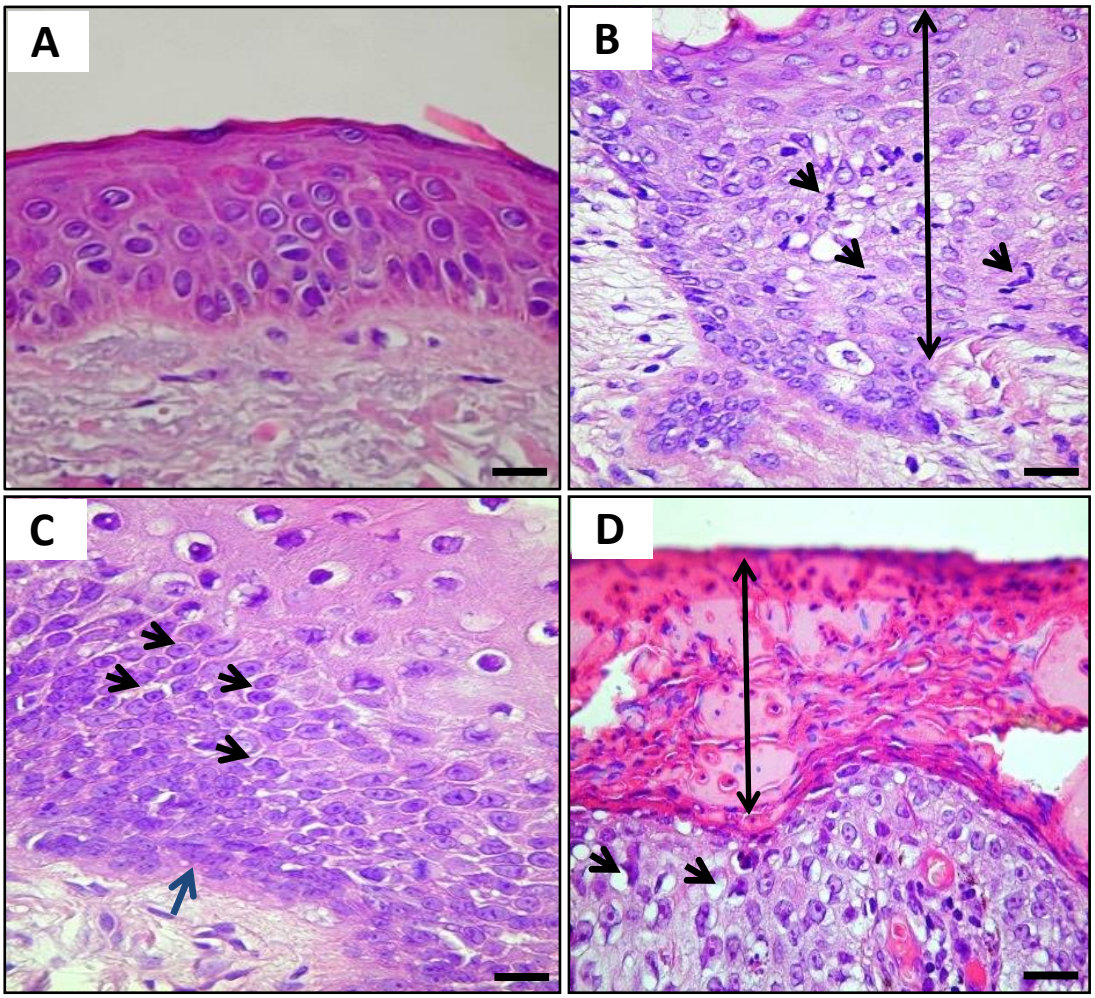


Figure 3.2 Dysplastic/non dysplastic features of AKs. **A:** 40x H&E image for normal skin. Dysplastic features in AKs are shown on images B-C and non-dysplastic features are shown in image D. **B:** 40x AK image showing mitotic figures (single headed arrows) and loss of polarity (disturbance in cellular orientation and loss of architecture and organization) (double headed arrow). **C:** 40x AK image showing nuclear atypia (black arrows), basal cell layer proliferation (blue arrow). **D:** 40x AK image showing other pathological features that are not specific for AK (non- dysplastic features); these features include vacuolization (single headed arrows), hyper/parakeratosis (double headed arrow). Scale bar = 10 µm.

The degree of dysplasia was scored according to the presence (1) or the absence (0) of any of the following parameters; loss of polarity, mitotic activity, basal cell layer proliferation, adnexal involvement with dysplastic cells and dyskeratosis, except for nuclear atypia which was scored as absent (0), mild (1), moderate (2). According to the severity of tissue damage, AKs were classified into 3 different grades of dysplasia as ((mild (scored 1-3), moderate (scored 4 -5) and severe (scored 6 -7)). The results are summarised in appendix table 8.2 and examples of the 3 grades of dysplasia are shown in figure 3.3. Out of the 69 AKs, 81.2% of the AKs showed a mild to moderate degree of dysplasia while 18.8% of AKs showed severe dysplasia. The result is shown in figure 3.4.

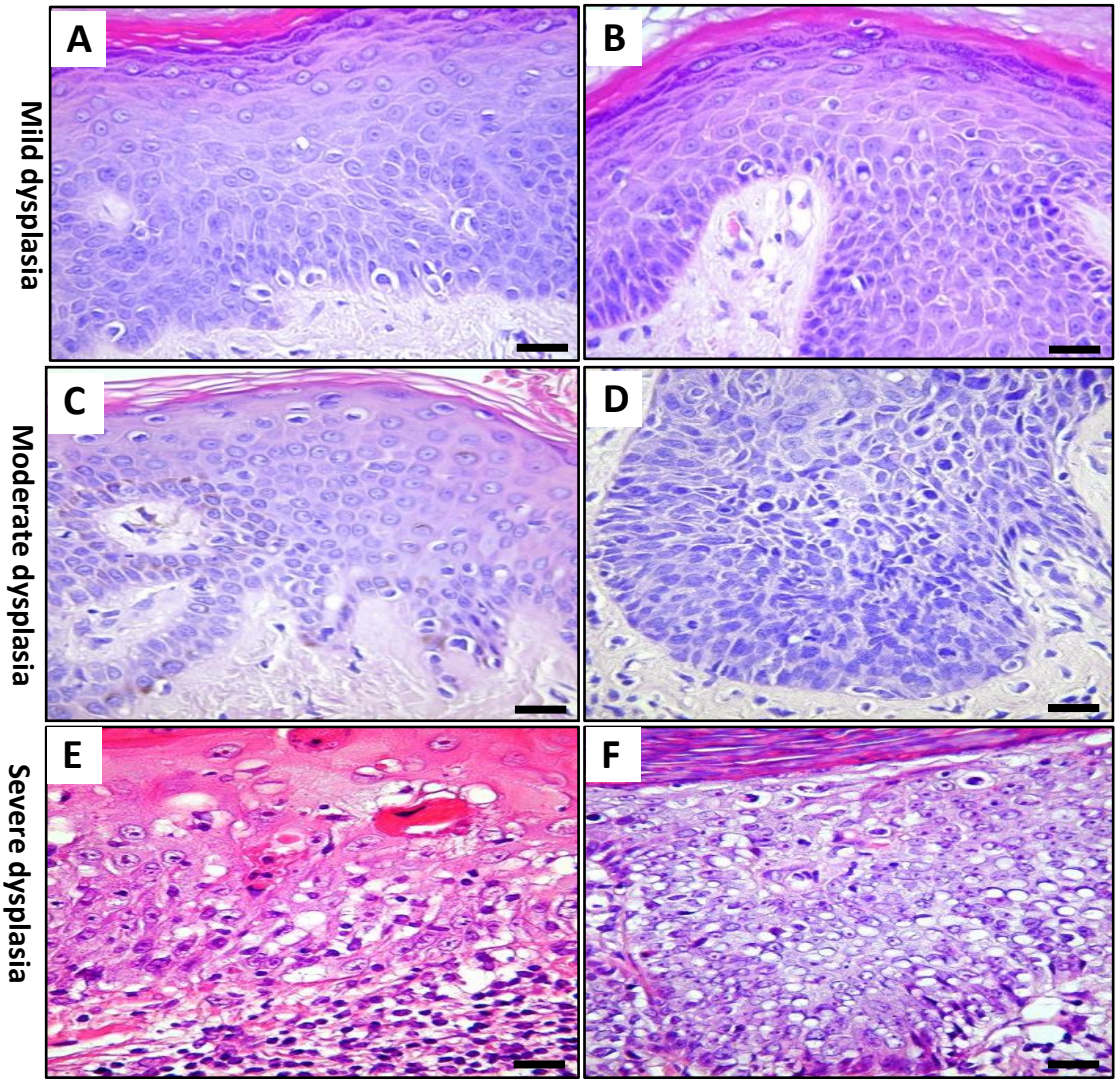


Figure 3.3 Dysplastic features of AKs. 6 AKs (2 examples for each grade of dysplasia) show the 3 different grades of dysplasia. **A&B:** Mild dysplasia (nuclear atypia and mitotic figures), **C&D:** Moderate dysplasia (nuclear atypia, mitotic figures and basal layer hyperproliferation), **E&F:** Severe dysplasia (nuclear atypia, mitotic figures, basal layer hyperproliferation and loss of polarity). Images taken at 40 fold magnification. Scale bar= 10 µm.

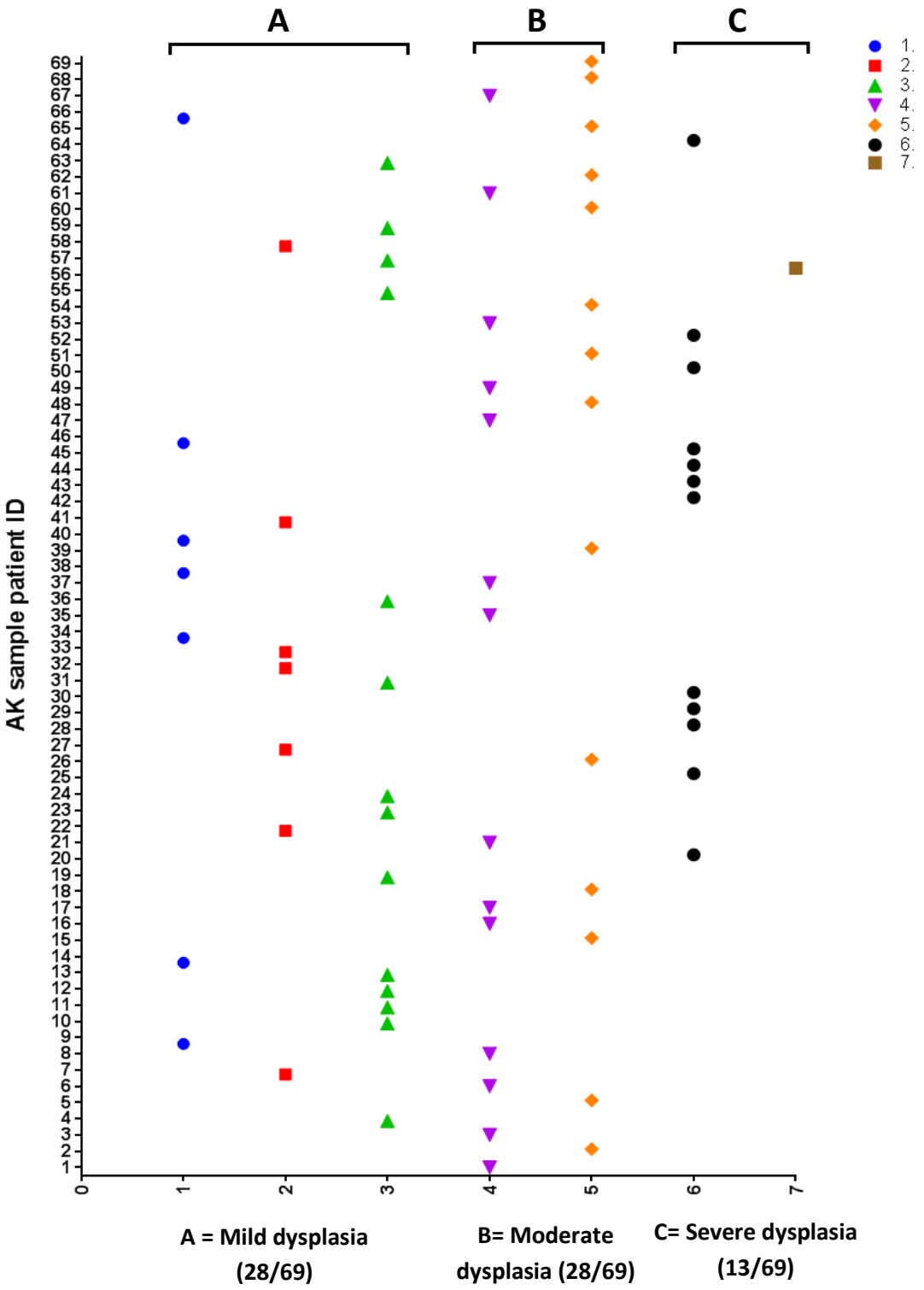


Figure 3.4 Dysplasia grades of AKs. Dysplasia was graded according to parameters which included nuclear atypia, loss of polarity, mitotic activity, basal cell layer proliferation, adnexal involvement with dysplastic cells and dyskeratosis. Each parameter was scored as present (1) or absent (0), except for nuclear atypia which was scored as absent (0), mild (1), moderate (2). The total score was then used to label the overall lesion as follows; 0 = no dysplasia. **A:** 1-3 = mild dysplasia. **B:** 4-5 = moderate dysplasia. **C:** 6-7 = severe dysplasia.

3.4.2.1.2 Counting dysplastic cells within actinic keratosis samples

The number of dysplastic cells within AK samples was quantified within 5 different images from each AK section using the 40x magnification H&E images and ImageJ software. The results are summarised in appendix table 8.3. Cells can be counted manually or automatically using the ImageJ package. While the manual method counts the cells based on the investigator's judgement, the automated method relies entirely on the ImageJ software (i.e. is a quicker counting method than the manual one). In the process of protocol optimisation, comparative assessment for the two methods was conducted. The process of counting and the use of ImageJ software are detailed in section 2.2.7. Different 40x images (N = 20) from 20 different AK sections were randomly selected and the numbers of dysplastic cells within the lesion were assessed using both manual and automated methods (detailed in section 2.2.7); examples of counts of dysplastic cells are shown in figure 3.5. Correlation coefficient analysis of the two variables (manual and automated) in relation to cell counting showed an *r* value equal to 0.98 with high correlation between the two methods as shown in figure 3.6.

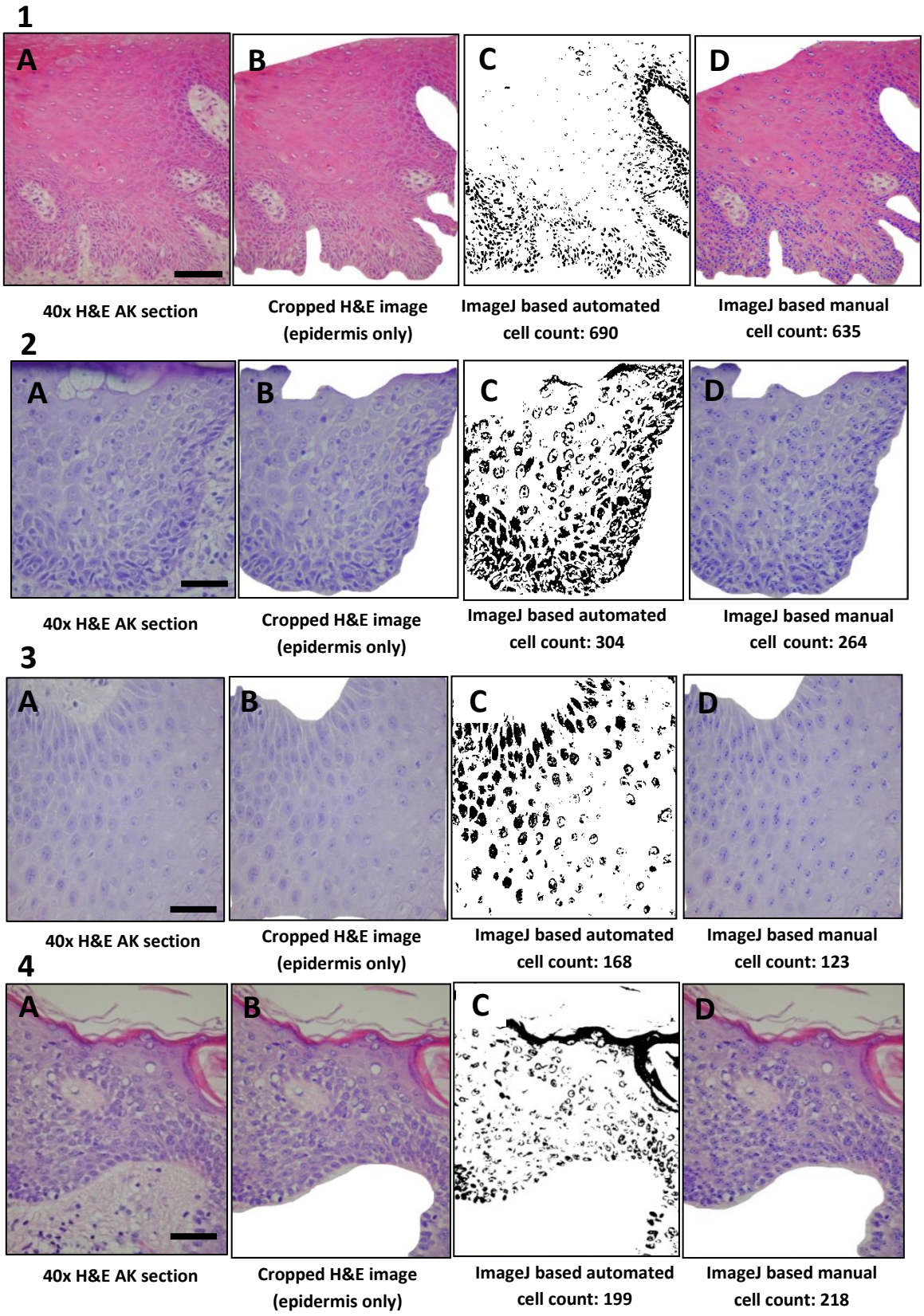


Figure 3.5 Automated and manual counting of dysplastic cells. 4 different AK samples (labelled as **1 -4**) are included. **A**: represents H&E image from each AK. **B:** represents image A with dermis excluded using ImageJ software (to ensure only epidermis is included within the counting). **C:** dysplastic cell counting using ImageJ analysis (automated), **D:** manual counting of dysplastic cells using the ImageJ software (to avoid re-counting of any cells). Images are taken at 40x magnification. Scale bar = 10 µm.

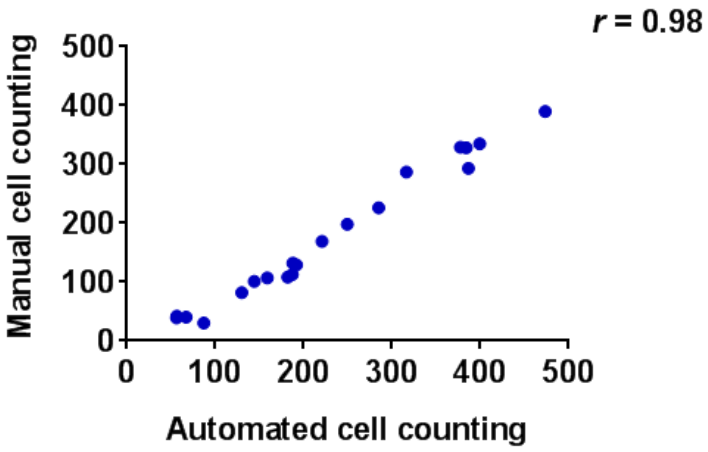


Figure 3.6 Correlation coefficient analyses between manual and automated counting methods of dysplastic cells. Dysplastic cells within 20 different AK 40x images were counted (one image per AK). *r* value = 0.98.

3.4.2.2 Immunohistochemistry staining

3.4.2.2.1 p53 staining

As discussed previously, anti-p53 antibody (DO-7) from Dako/UK was used to detect the nuclear p53 protein staining within epidermal keratinocytes. Examples of p53 staining are shown in figure 3.7.

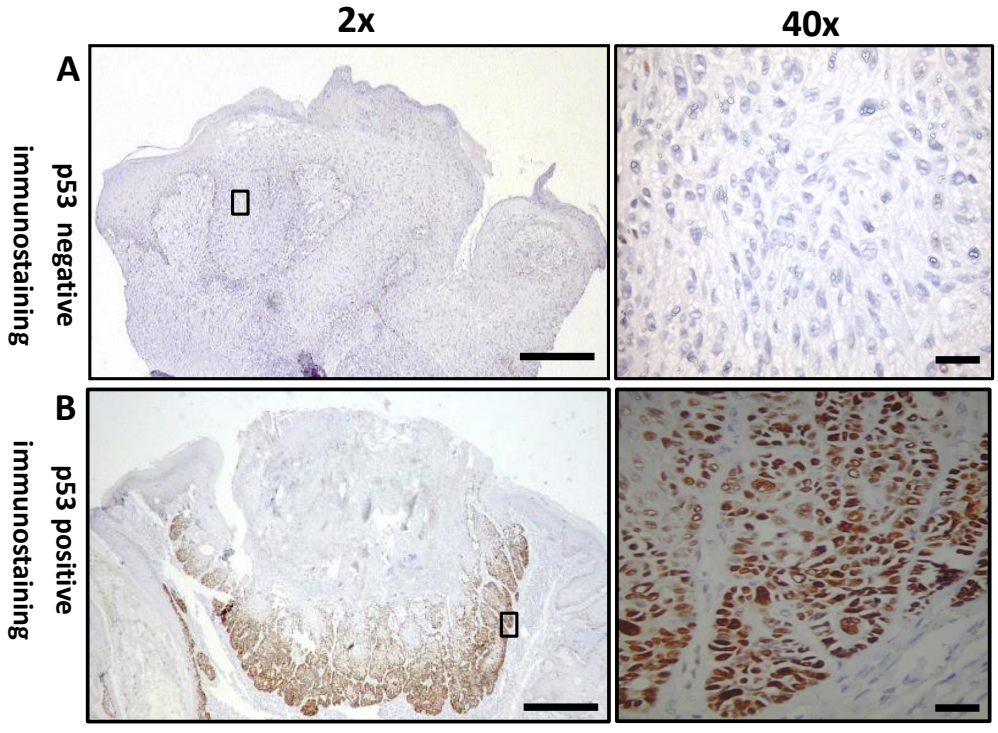


Figure 3.7 p53 staining within AKs. **A:** A representative AK section with negative p53 staining. **B:** A representative AK section with positive p53 nuclear staining. Left images 2x and right images 40x magnification. Scale bar; 500 & 10 µm for 2x and 40x magnifications respectively.

The number of epidermal p53 positive nuclei was counted within 5 representative epidermal 40x images from each AKs using ImageJ software. The images were selected to cover the dysplastic areas within the AK sections; a unit area equal to 0.056 mm^2 was selected. p53 staining was positive in 87% of the AK samples and was predominantly seen in the dysplastic area of the lesion (when compared with the H&E section). Examples on the process of ImageJ analysis of p53 staining is summarised in figure 3.8 and the complete results are summarised in appendix table 8.4.

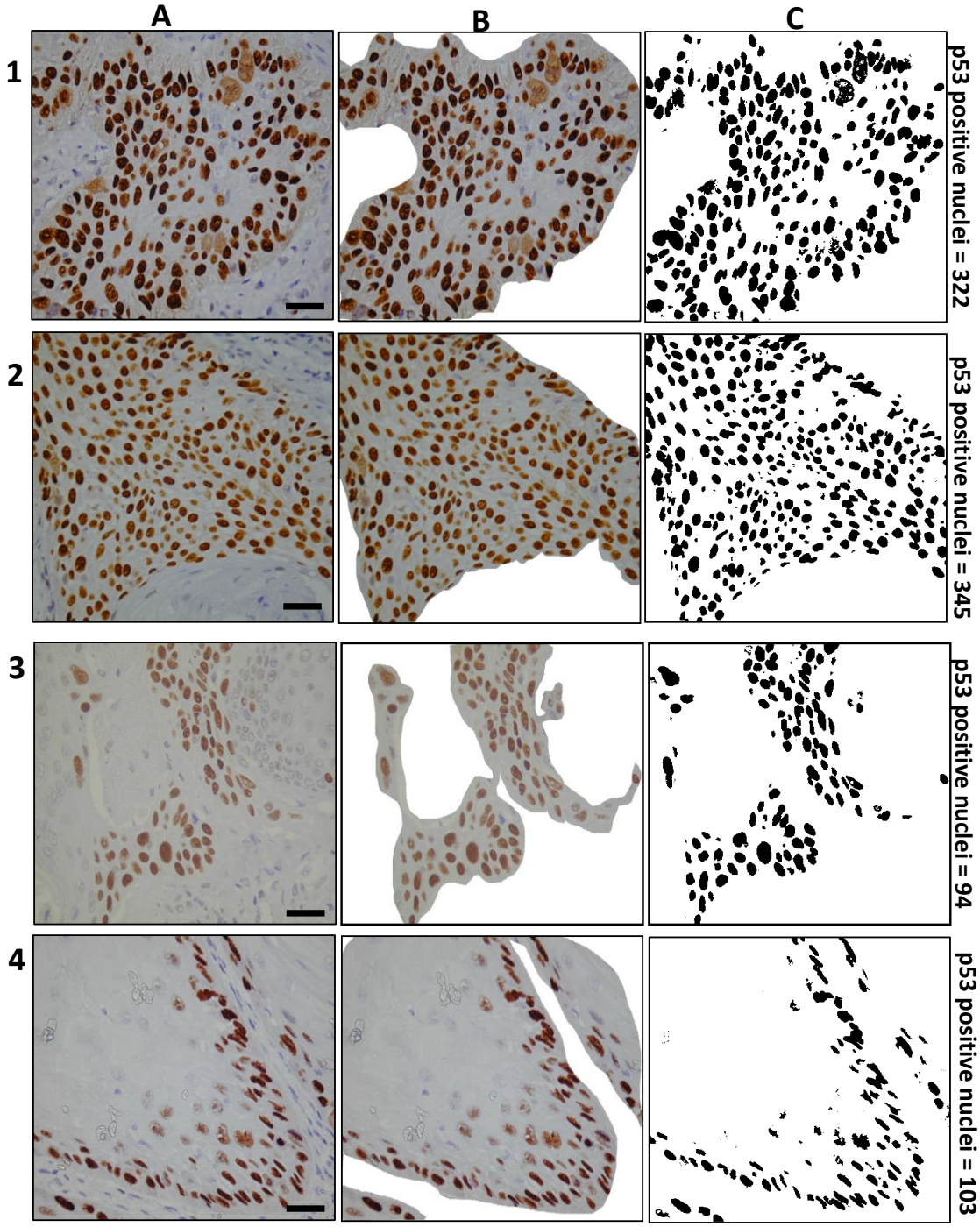


Figure 3.8 ImageJ based counting for p53 staining. 4 AK samples (rows **1** -**4**) are shown in this figure, with numbers of p53 positive nuclei counted in the section on the right. Column **A:** 40x magnification of p53 stained AK image. Column **B:** p53 positive component of image A, where non-epidermal tissue was excluded using the crop facility on ImageJ software. Column **C:** The images were processed in ImageJ and then converted into a binary image (black and white). p53 positive nuclei were quantified by ImageJ using 'analyse particles' option on the software and only particles that are larger than 50 pixels (to ensure that these represented p53 stained nuclei) were included in counting. A standardised unit of section area equal to 0.056 mm^2 was used to analyse the staining. Scale bar = 10 µm.

In addition, the relationship between p53 staining and grade of dysplasia within AKs was also analysed and summarised in figure 3.9 and appendix table 8.5. The aim of these analyses was to investigate if there was preferential expression of p53 protein within any of the groups of dysplasia. Statistical analysis using one way ANOVA showed no significant difference between the three grades of dysplasia in relation to the percentage of p53 staining within AKs (p = 0.54).

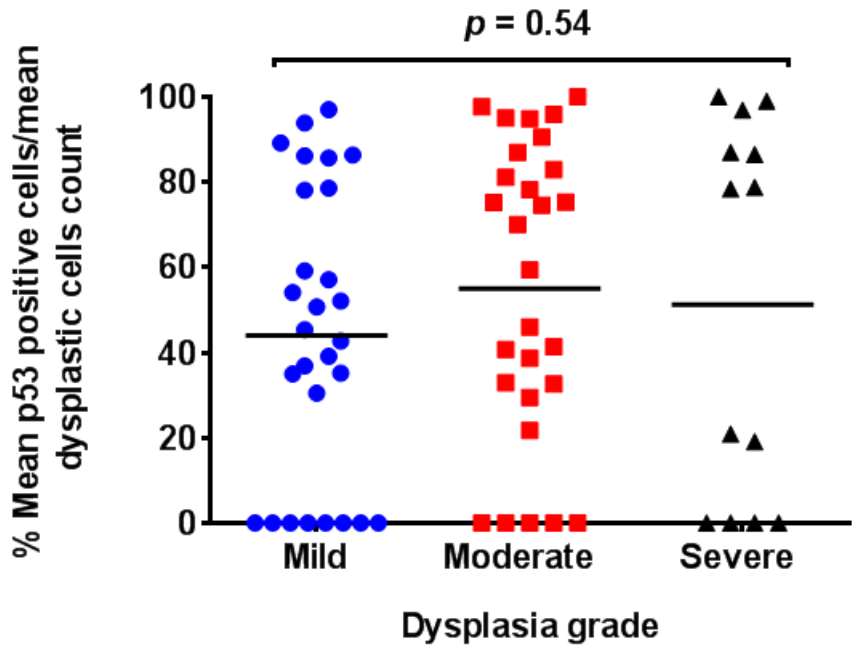


Figure 3.9 The percentage of the mean p53 positive cells to dysplastic cells over the grades of dysplasia (mild, moderate and severe) of AKs. There was no statistical difference between the 3 grades of dysplasia in relation to the percentage of p53 staining within AK samples using one way ANOVA (p = 0.54).

3.4.2.2.2 beta catenin staining

The normal localisation of beta catenin protein within the cell is the cell membrane. Abnormal localisation of beta catenin protein within different tissue, including skin, has been reported (Brasanac et al., 2005). Within the group of AKs in this study, the staining of beta catenin protein was variable in terms of the cellular localisation of the protein (normal membranous and aberrant (cytoplasmic, nuclear)) and the intensity of the membranous staining within the sections (normal, reduced or lost). The results of beta catenin staining were very heterogeneous even within the same AK section which made it difficult to use ImageJ software in the analysis. Consequently, manual estimations of the intensity and the cellular localisation of the protein were performed. 29.68% AK sections showed homogenous membranous staining and 71.87% showed heterogeneous staining. The results are summarised in appendix table 8.6 and examples of the staining are shown in figure 3.10. Overall, it was considered that beta catenin staining was not helpful in delineating the dysplastic area in AKs.

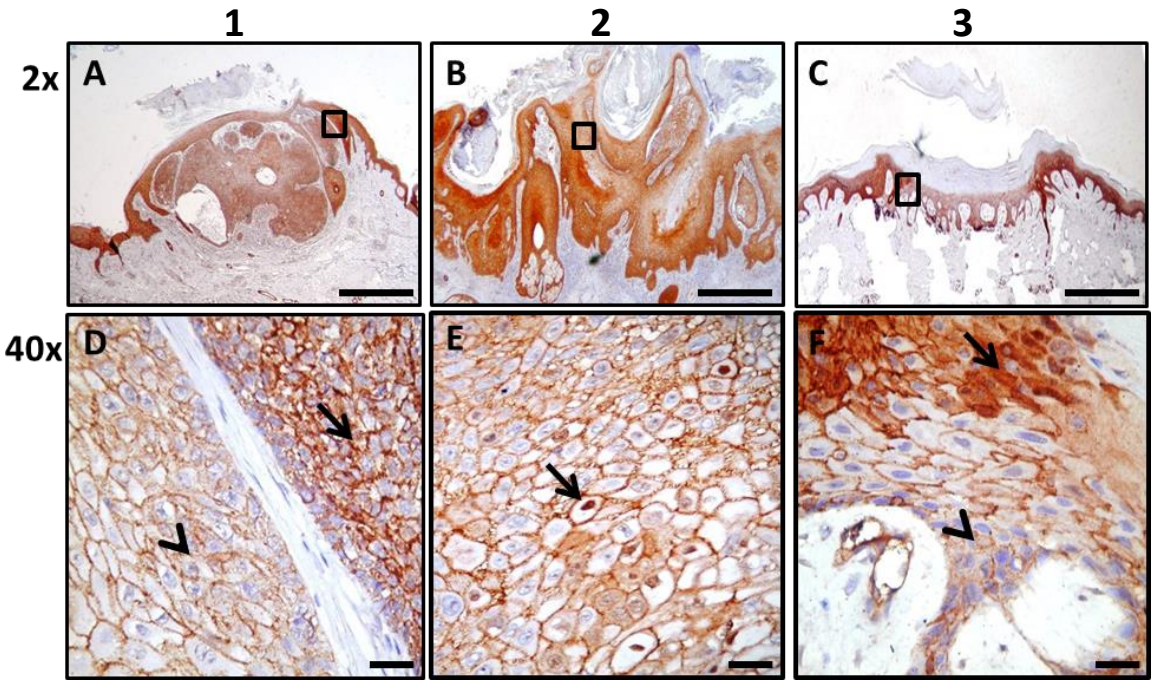


Figure 3.10 beta catenin staining within AKs. Different intensities or patterns of beta catenin staining within the same section were often seen. 3 different examples (**1, 2** and **3**) are shown. The first row show 2x images (**A, B** and **C**)with black boxes that are represented in more details in 40x magnification images on the second row (**D, E** and **F**) (the variety of staining patterns is seen best on the 40x images). **D:** Represents example of normal membranous beta catenin staining (single headed arrow) and area of reduced intensity of the normal membranous staining (arrow head), **E:** Represents example of beta catenin nuclear (single headed arrow) staining, **F:** Represents example of beta catenin staining showing area of cytoplasmic localisation (single headed arrow) and loss of membranous staining (arrow head). Scale bar; 500 & 10 µm for 2x and 40x magnifications respectively.

3.4.2.2.3 Staining for dermal immune infiltrate

Many AKs exhibited an immune infiltrate within the dermis on H&E slides which could possibly interfere with delineating the AK dysplastic keratinocytes by an inadequately experienced researcher. Thus, it was important to be able to consistently distinguish dysplastic epidermal keratinocytes from perilesional dermal immune cells to ensure the dissection of pure lesional epidermis in downstream experiments. Thus, dermal immune cells were recognised using antibodies that target different immune cell antigens.

3.4.2.2.3.1 CD4 staining

CD4 is a T cell antigen that can be visualised on the cytoplasmic membrane of T cells upon immunostaining with a specific anti-CD4 antibody. CD4 staining results for AKs were analysed and counted using ImageJ software; the results are summarised in appendix table 8.7 and examples of the staining are shown in figure 3.11.

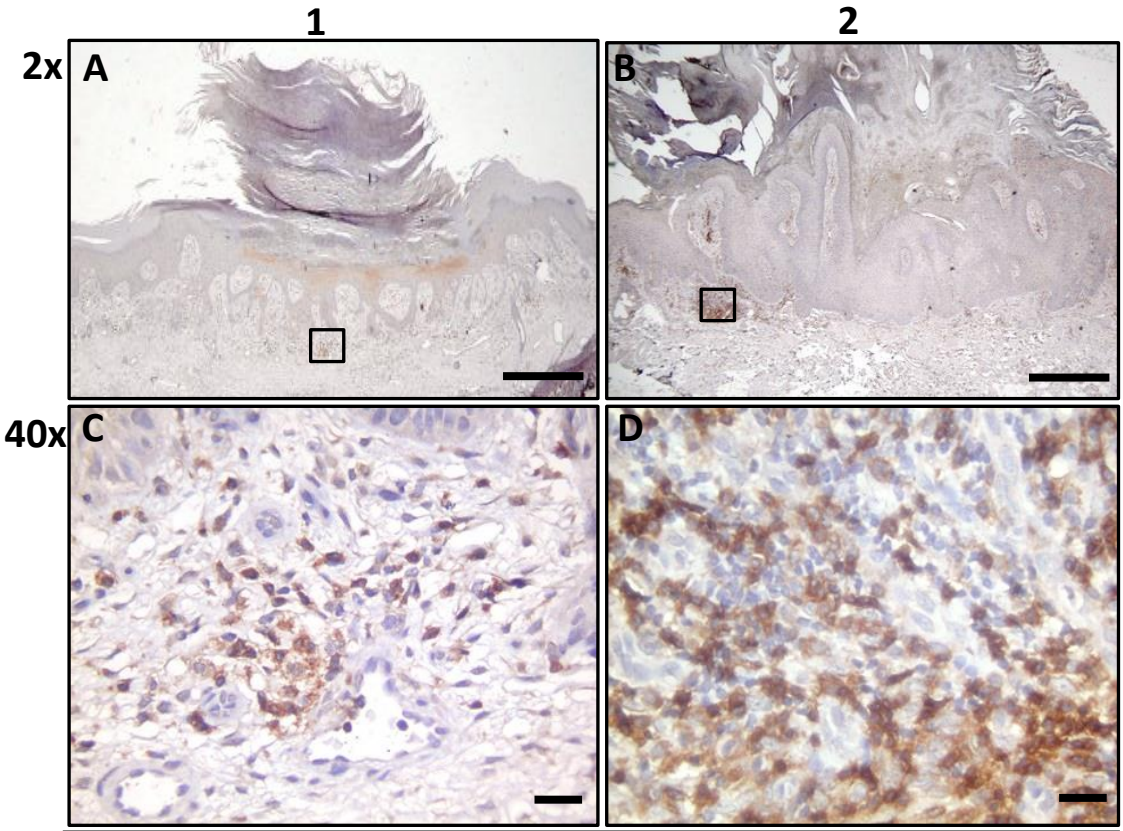


Figure 3.11 Immunostaining of CD4 antigens within the dermal immune infiltrate of AK lesions. Positive staining is seen as brown membranous colouration of the dermal immune cells and the remaining perilesional cells are blue. **1&2:** are examples of CD4 membranous staining within two different AK sections. **A & B:** images represent 2x magnification while **C & D:** images represent 40 x magnifications. The boxes on the 2x images represent the areas shown on the **C & D** images. Scale bar; 500 & 10 µm for 2x and 40x magnifications respectively.

Statistical analysis using one way ANOVA was used to compare the proportion of CD4 + T-cells within the perilesional immune infiltrate with the grade of dysplasia (mild, moderate and severe) in AKs as shown in figure 3.12. The mean number of CD4+ T-cells present in the perilesional infiltrate was found to be 33.7%, 31.8% and 37.8% within the mild, moderate and severe dysplasia grade groups respectively and there was no statistical difference in the number of the CD4+ T cells between the three groups (p = 0.62).

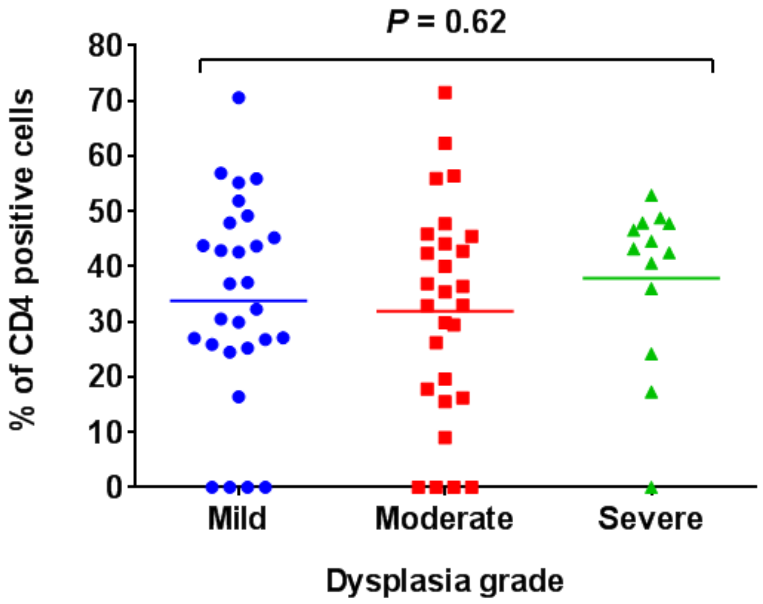


Figure 3.12 The percentage of CD4 staining in relation to the grade of dysplasia (mild, moderate and severe) within AKs. Results are expressed as a percentage of the cells which are CD4+ in the perilesional immune cell infiltrate using mean values calculated in appendix table 8.7. There was no statistical difference between the three groups of dysplasia in relation to the CD4 staining within AKs using one way ANOVA) (p = 0.62).

3.4.2.2.3.2 CD8 staining

Using ImageJ, the membranous staining of CD8 on perilesional dermal T-cells was analysed; the results are summarised in appendix table 8.8 and examples of the staining are shown in figure 3.13.

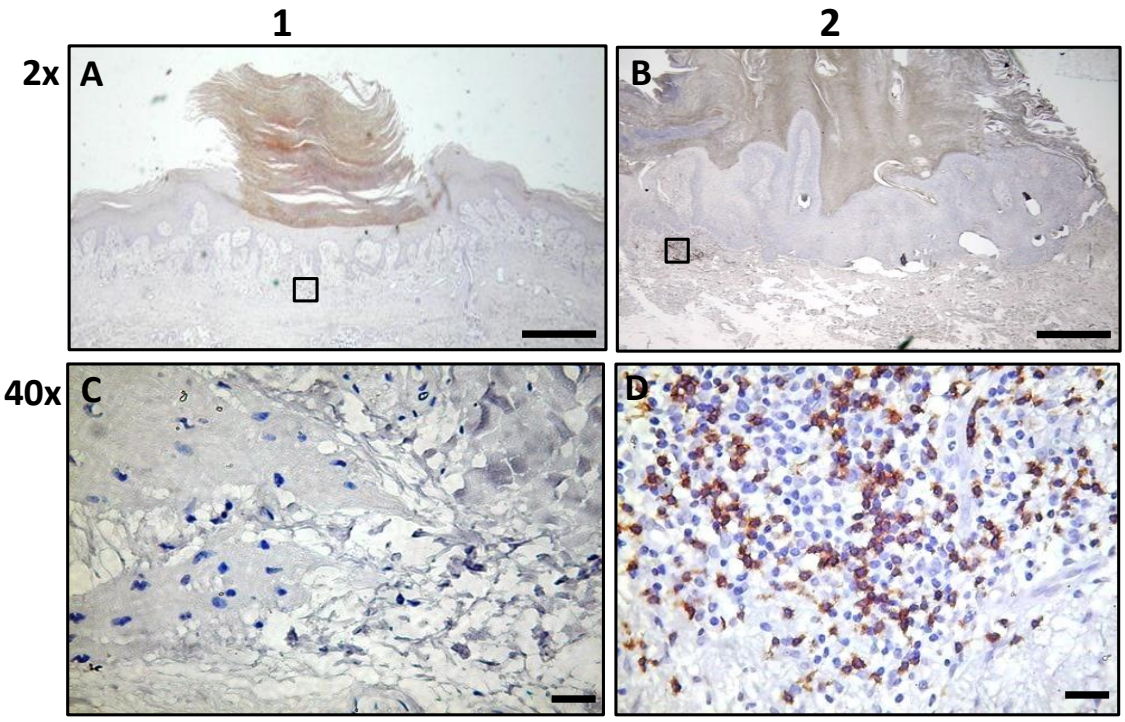


Figure 3.13 CD8 staining within the dermal perilesional immune infiltrate in AKs. The brown membranous staining represents the CD8+ T cells and the remaining perilesional cells are blue. **1&2:** Two different examples. **A & B:** images represent 2x magnification while **C & D:** images represent 40x magnification. The boxes on the 2x images represent the areas shown in the **C & D** photographs. Scale bar; 500 & 10 µm for 2x and 40x magnifications respectively.

One way ANOVA was performed to compare the distribution of CD8+ T-cells with the grades of dysplasia (mild, moderate and severe) as shown in figure 3.14. The mean number of CD8+ T-cells in the perilesional infiltrate was 35.5%, 37.6% and 44.3% within the mild, moderate and severe dysplasia groups respectively and there was no statistical difference in the numbers of CD8+ cells between the three groups (p = 0.92).

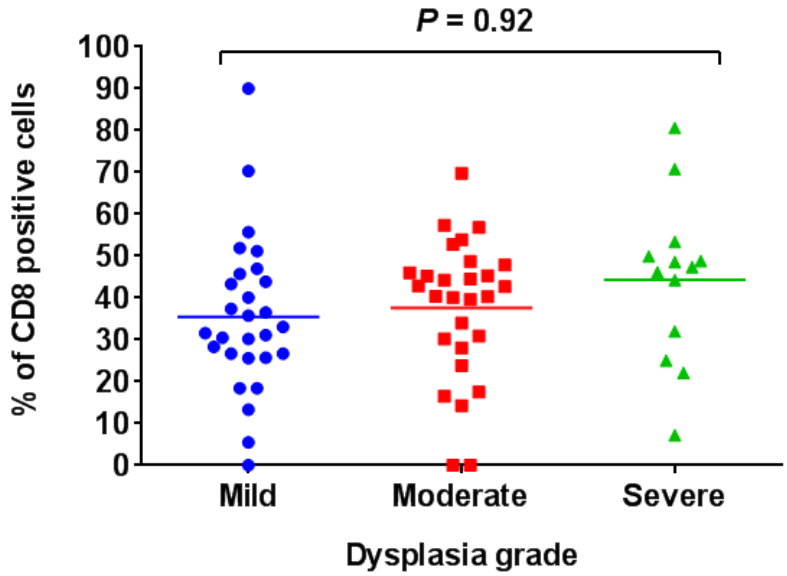


Figure 3.14 The percentage CD8 positive staining cells in the perilesional immune cell infiltrate in relation to grade of dysplasia within AKs. Results are expressed as a percentage of the perilesional immune cells which are CD8+ using mean values calculated in appendix table 8.8. There was no statistical difference between the three groups of dysplasia in relation to the CD8 staining within AKs (p = 0.92 using one way ANOVA).

3.4.2.2.3.3 FOXP3 staining

FOXP3 staining of T cells, which identifies regulatory T cells (Tregs), was assessed using ImageJ software within the perilesional immune infiltrate; the results are summarised in appendix table 8.9 and examples of the staining are shown in figure 3.15.

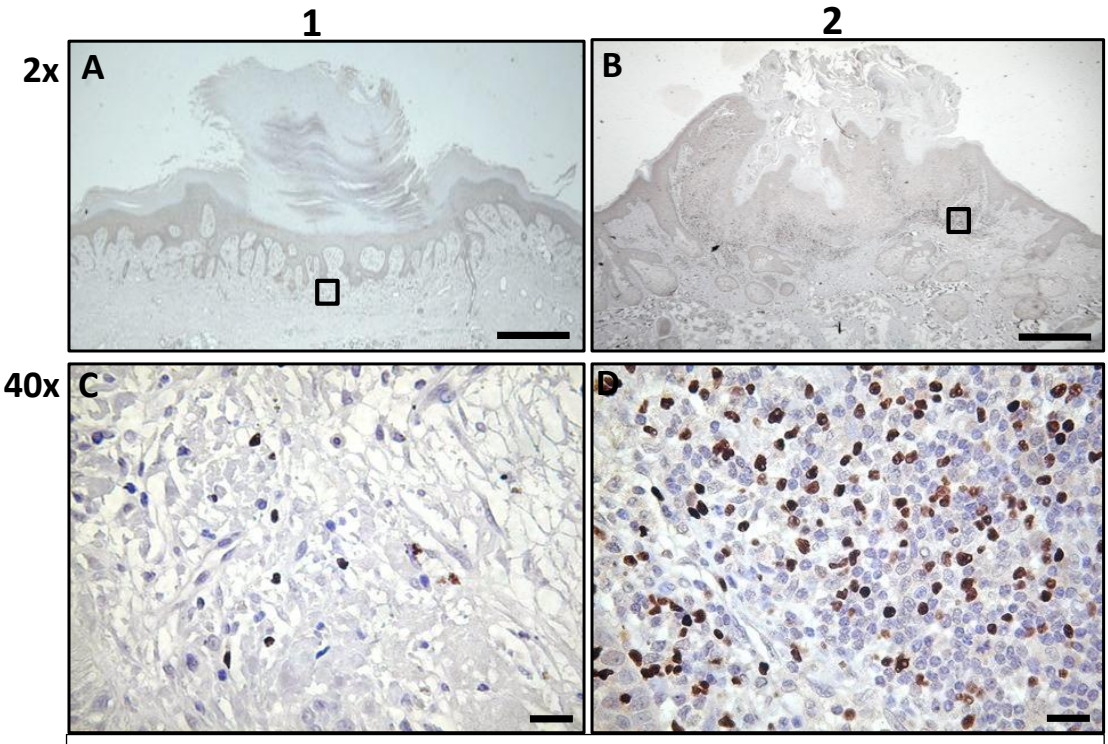


Figure 3.15 FOXP3 staining within AKs. The nuclear brown colouration represents the FOXP3 positive cells and the remaining blue cells within the perilesional T cells are FOXP3 negative. **1&2:** Examples of FOXP3 staining in two AK samples. **A&B** images were taken at 2x magnification whereas **C&D** represent 40x magnification images. The boxes in **A&B** represent the areas shown in the **C&D** images. Scale bar; 500 & 10 µm for 2x and 40x magnifications respectively.

The mean number of FOXP3+ T-cells within the perilesional infiltrate was 21.7%, 24.1% and 26.4 within the mild, moderate and severe dysplasia groups respectively and there was no statistical difference in the distribution of the FOXP3+ cell between the three groups (p = 0.67). The result is shown in figure 3.16.

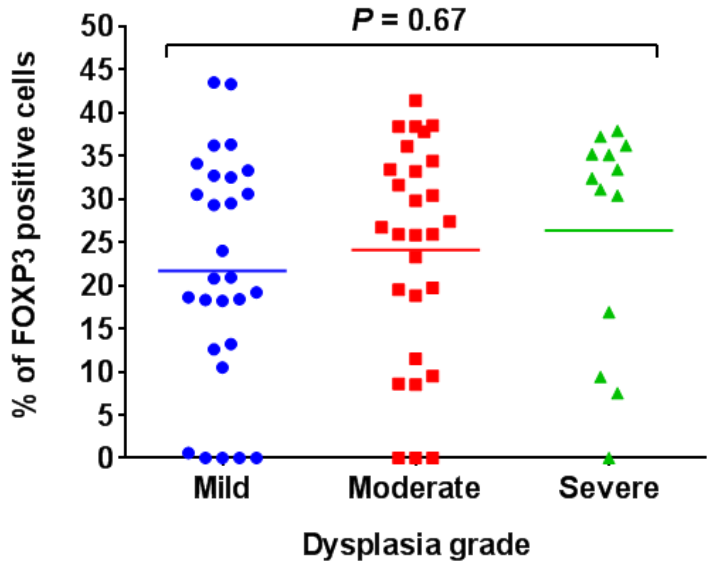


Figure 3.16 The percentage of FOXP3+ T cells in the perilesional infiltrate in relation to the degree of dysplasia within AKs. Results are expressed as a percentage of cells within the perilesional immune infiltrate which are FOXP3+ using mean values calculated in appendix table 8.9. There was no statistical difference between the three groups of dysplasia in relation to the FOXP3 staining within AKs (p = 0.67 using one way ANOVA).

3.4.3 Laser capture microdissection

3.4.3.1 Selection of AK samples for WES

The next step after acquiring the ability to confidently identify the dysplasia within AKs was sample selection for WES analysis. The main selection criterion used was the size of the dysplastic area visible on the AK tissue; this is because sections with larger areas of dysplasia were likely to provide more DNA for WES. The aim was to get sufficient yield of DNA that met the requirement for WES without the need to conduct whole genomic amplification. A previous study showed that up to 60 sections, each 8 µm thick, from cSCC yielded sufficient DNA for WES analysis (South et al., 2014). However, other factors also influenced the decision about which AKs to choose for WES. These included; whether immune cells had infiltrated into the dysplastic area (thus making microdissection difficult), the library preparation kit (different library preparation kits require different DNA amounts) and the nature of the tissue (FFPE). In addition, the high cost of WES also influenced the decision, in terms of the number of AKs considered for WES. 5 samples from a total of 69 AKs were therefore selected for WES.

The selected samples were predominantly located on the skin of the head and neck region. There were 3 males and 2 females, with a median age of 83 years. The grade of dysplasia within the AK samples ranged from 5 to 6. Samples showed some variation in the different immunostaining positivity / patterns within the lesions. The patient's details is summarised in table 3.2.

Table 3.2 Details of AK samples selected for WES (N = 5). The median age of the patients was 83 and the female: male ratio was 2:3. All lesions were from sun exposed skin. The percentage of the dysplastic area of the lesions positive for (p53) and immune cell infiltrate (CD4, CD8, FOXP3) staining markers are shown, whereas for beta catenin the pattern of staining is provided. The grade of dysplasia (and dysplasia score) indicates that the AKs displayed either moderate or severe dysplasia. M = Male and F = Female.

Patient ID	Age at biopsy	Gender	Site of the lesion	p53%	beta catenin	CD4%	CD8%	FOXP3%	Dysplasia grade
28	74	M	Right temple	0	Reduced membranous	17.3	70.8	35.2	Severe (+6)
30	85	M	Left cheek	9.34	Reduced membranous	47.9	22.1	33.4	Severe (+6)
51	86	F	Left jaw line	36.9	~95% membranous ~5% nuclear	36.9	44.5	36.1	Moderate (+5)
54	62	M	Left cheek	20.7	~60% membranous ~30% loss of staining ~10% nuclear	44.1	56.8	31.6	Moderate (+5)
68	83	F	Left middle finger	95.9	Normal membranous	56.4	44.2	30.4	Moderate (+5)

3.4.3.2 LCM results

The selected AK blocks were cut and stained as outlined in section 2.2.8. Both (H&E) and cresyl violet acetate staining were recommended to visualise the sections by the LCM Leica microsystem protocol as per the sample preparation for LCM in the Leica microsystem handbook manual. However, according to Cummings et al. (2011) cresyl violet acetate is superior to H&E staining in terms of preserving the integrity of extracted DNA. In addition, the cresyl violet staining protocol requires less time than H&E staining (1-2 minutes for cresyl violet staining verses 15 minutes for H&E staining). Thus, cresyl violet acetate was selected to stain the FFPE AK sections prior to LCM in this study.
The target areas (dysplastic and adjacent normal looking skin) were then dissected using the LCM microscope as described in section 2.2.8. However, during the use of the LCM microscope, it was noted that the laser beam failed to cut through the hyperkeratotic part of the AK adequately, therefore the stratum corneum was removed manually from

hyperkeratotic AK lesions using straight dissecting needle (Thomas Scientific). The results are shown in figure 3.17 and 3.18.

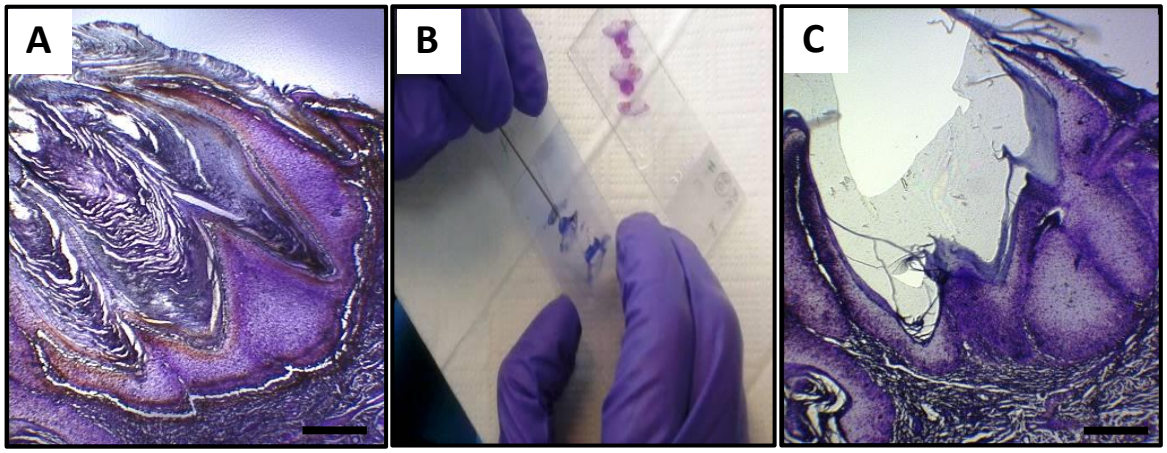


Figure 3.17 The process of manual removal of stratum corneum from 10µm thick AK sections. **A:** represents a 4x AK image stained with cresyl violet acetate before stratum corneum removal. **B:** manual removal of the stratum corneum from a stained AK section with a straight dissecting needle (Thomas Scientific), with H&E slide used for assistance in recognising the stratum corneum. **C:** An example of an AK sections following manual stratum corneum removal. Scale bar = 200µm.

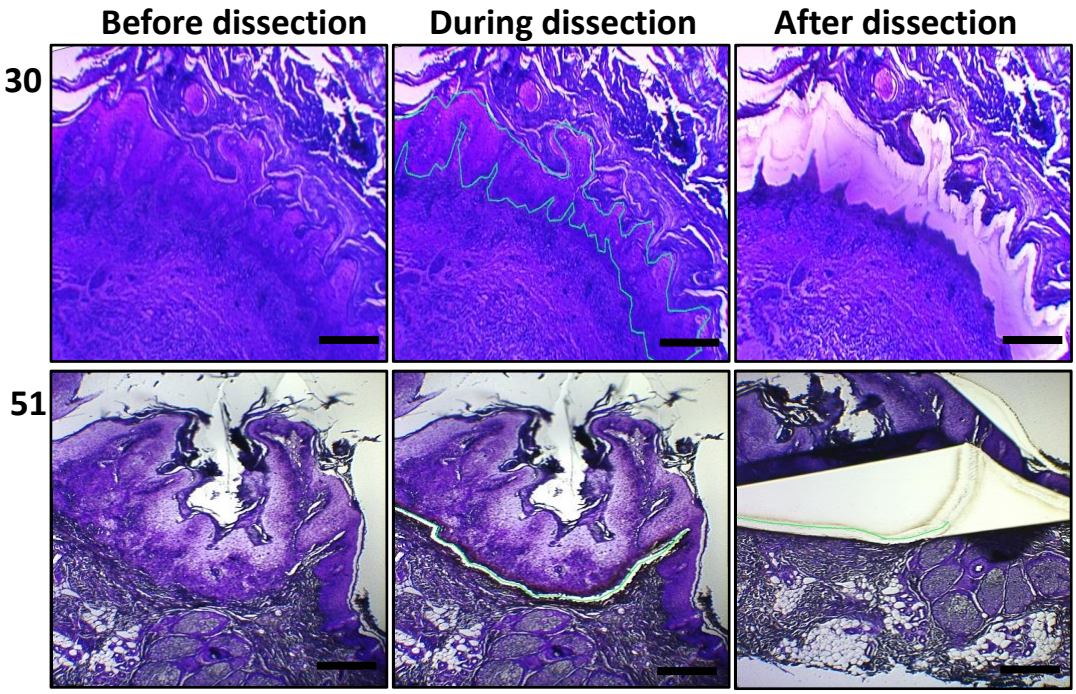


Figure 3.18 The process of dysplastic tissue dissection by LCM. **30 & 51:** Two examples of AKs (subjects 30 and 51) selected for WES are shown in this figure. Both samples were hyperkeratotic. The dysplastic area in sample 30 sections was dissected without the manual removal of stratum corneum, in contrast to sample 51 where the section was dissected after removal of stratum corneum. The first column shows sections before dissection under the LCM. The second column illustrates the target (dysplastic) area demarcated within the green line. The third column shows the area during and after it has been dissected using LCM. The dissected tissue falls into the collection microcentrifuge tubes. Scale bar = 200µm.

Adjacent normal looking skin was also removed from the same tissue sections and collected in a separate tube so that the extracted DNA from this could be used as a reference for variant calls within the dysplastic area of the AKs. In order to reduce potential contamination between the two targeted areas (dysplastic and normal looking skin), a sufficiently wide area of skin between the two targeted tissue areas was dissected before removal of the adjacent normal looking skin into the separate microcentrifuge tubes. The results are shown in figure 3.19.

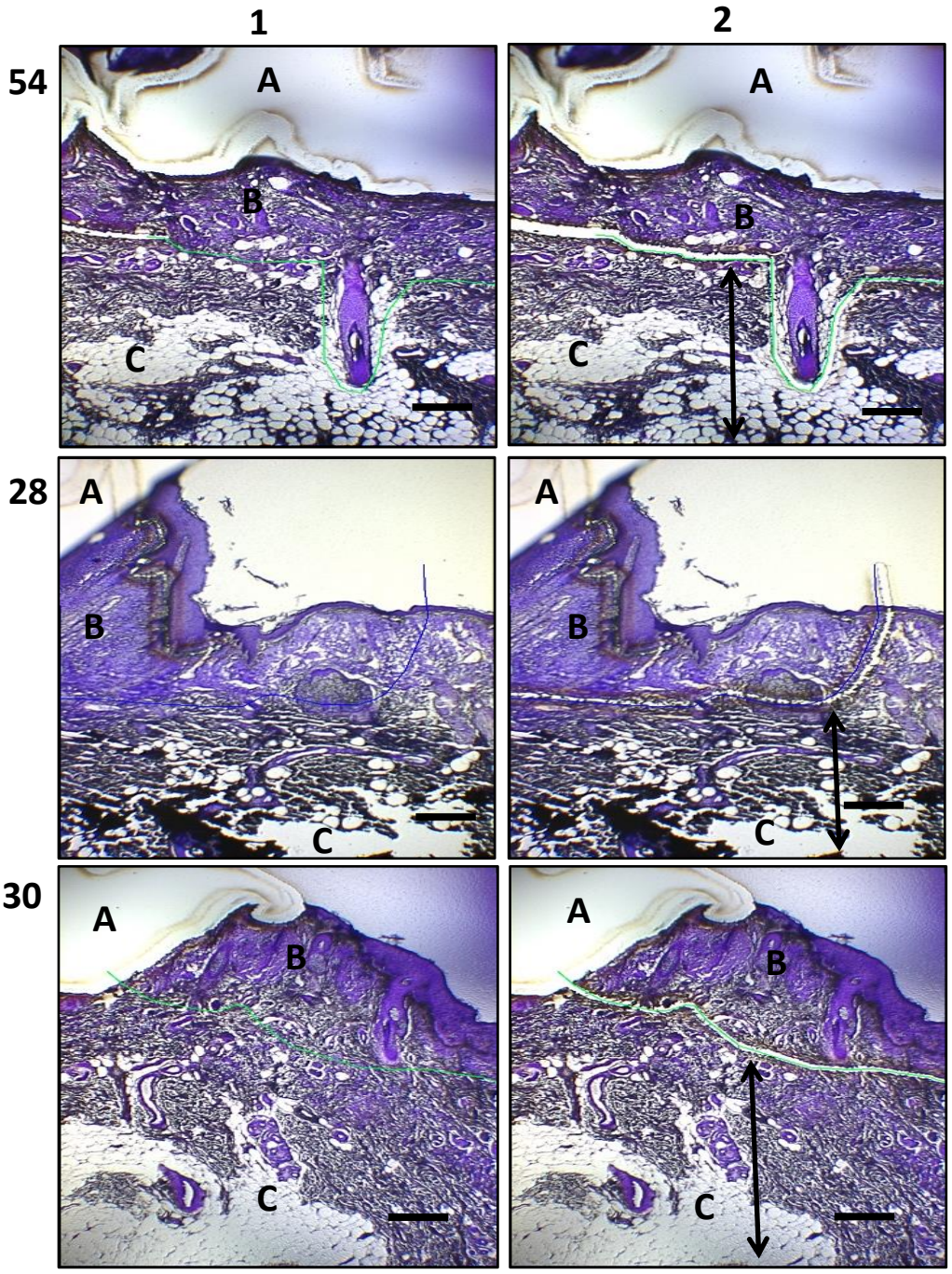


Figure 3.19 Normal control skin adjacent to the selected AKs for WES. 3 examples are shown (AK samples 54, 28 and 30). **Column 1:** shows the demarcation line (green or blue). **Column 2:** shows the same sections after dissection. **A:** the site of the dissected dysplastic tissue, **B:** the border region between normal and abnormal areas and **C:** the tissue that represents the normal control skin (indicated by the double headed arrows on column 2). Scale bar = 200µm.

3.4.4 Isolation of genomic DNA

3.4.4.1 Optimisation of DNA extraction, quality and quantity assessment procedures

FFPE samples are considered to offer a valuable resource for DNA studies (Sah et al., 2013, Xuan et al., 2013). Although the quality of DNA can be affected due to the cross-linking bonds introduced by formalin fixation (Xuan et al., 2013), good quality DNA can be obtained from FFPE in a number of cases as seen in previous work on other FFPE tissue types (Van Allen et al., 2014). In the current study, the QIAamp DNA FFPE Tissue Kit from Qiagen was used for DNA extraction as it was likely to produce DNA of a quality suitable for WES.

3.4.4.1.1 DNA quantity assessment

As per previous work by Simbolo et al. (2013) and the recommendation from the user manual of the Agilent SureSelect V5 kit (the library preparation kit that was subsequently used for WES in this study), the Qubit 2.0 fluorometer was used to determine the quantity of extracted DNA. This technique is based on the use of a fluorescence dye that binds specifically to double strand DNA molecules, thus allowing accurate and specific quantification of the amount of DNA. In addition, the Qubit 2.0 fluorometer is considered superior to Nanodrop spectrophotometry as the latter can overestimate the actual DNA concentration by up to 15 fold (Sah et al., 2013).

3.4.4.1.1.1 Assessment of DNA yield per the number of targeted sections

In order to estimate the number of AK tissue sections that would be required to obtain the target DNA concentration (30 - 50 ng/µl, total 1-3 µg) for WES, a comparative experiment that compared the DNA yield from different number of FFPE sections (in this case of normal skin) and correlating the results to the size and cellularity of the target tissue was conducted. Human skin tissue was obtained from redundant breast skin which had been taken during mastectomy of patients with breast cancer at Princess Anne Hospital. Skin pieces (N = 5) were obtained from 5 different patients and then were formalin fixed and paraffin embedded. 10µm sections were cut from each block and dewaxed in xylene and rehydrated in graded ethanol. The sections then were scraped off the slides and collected in 1.5 ml microcentrifuge tubes containing 180 µl of ALT lysis buffer from Qiagen and the QIAamp DNA FFPE Tissue Kit was used for DNA extraction as per manufacture protocol in section 2.2.9. As the main cellular area in the skin is the

epidermis, the size and the number of epidermal cells were measured and counted respectively on H&E stained slides as shown in table 3.3.

Table 3.3 Information on the breast skin used in the comparison of the yield of DNA per number of 10µm tissue sections from the same sample (N=5). The size of the skin tissue was kept consistent at ~ 6 x 5 mm². ImageJ analysis of H&E stained sections, magnified 2x, was used for measuring the size of the epidermis in mm². The number of keratinocytes within 5 H&E images (at 40x magnification) for each section was counted. For result consistency, keratinocytes were counted within a unit of section area equal to 0.056 mm² and the mean number of cells counted within five different areas across the section was calculated.

Sample ID	Size of the section in mm²	Size of epidermis in mm²	Mean number of cells within 0.056 mm² of five representative (40x) images of the epidermis	The number of keratinocytes within the whole epidermis
1	~ 6 X 5	2.43	176	427.68
2	~ 6 X 5	2.32	267	619.44
3	~ 6 X 5	2.08	184	382.72
4	~ 6 X 5	2.14	182	389.48
5	~ 6 X 5	2.89	212	612.68

DNA was extracted using different numbers of 10µm FFPE breast skin sections (i.e. 10, 20, 30 and 40 tissue sections) and quantified by Qubit 2.0 fluorometer. Correlation coefficient analysis showed a high correlation between the number of sections and DNA yield (*r* value range between 0.93 – 0.98). Based on these results, it was concluded that 40 sections of AKs would be the minimum number of sections that could be used (taking into consideration the size of the dysplastic area and number of cells within the dysplastic area). The results are summarised in table 3.4 and figure 3.20.

Table 3.4 Summary of DNA yield per number of sections used for DNA extraction from 5 FFPE breast skin samples. DNA was extracted from 10, 20, 30 and 40 skin sections; each 10µm thick, for each sample and the resulting DNA was quantified by Qubit 2.0 fluorometer.

Sample ID	Number of sections	Number of cells/number of sections	DNA yield in ng/µl
1	10	4276.8	4.4
	20	8553.6	7.5
	30	12830.4	11.2
	40	17107.2	15.7
2	10	6194.4	6.1
	20	12388.8	11.6
	30	18583.2	13.5
	40	24777.6	16.8
3	10	3827.2	5.7
	20	7654.4	9.6
	30	11481.6	11.3
	40	15308.8	15.5
4	10	3894.8	3.5
	20	7789.6	6.9
	30	11684.4	10.4
	40	15579.2	16.2
5	10	6126.8	4.9
	20	12253.6	7.2
	30	18380.4	14.1
	40	24507.2	17.9

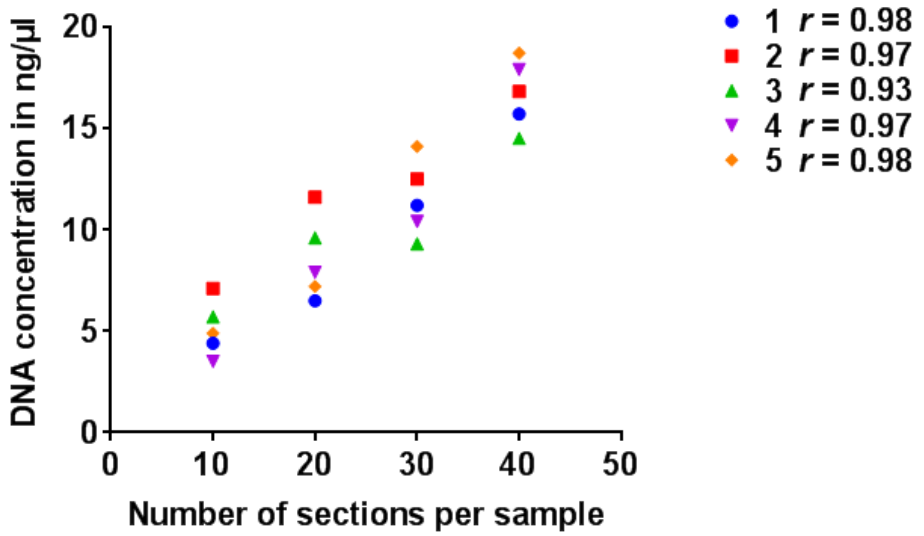


Figure 3.20 Comparison of DNA yield using different number of sections per skin sample. 10, 20, 30 and 40 sections, each 10µm thick, from 5 FFPE breast skin samples were put into separate microcentrifuge tubes and DNA extracted using a QIAamp DNA FFPE Tissue Kit. Qubit 2.0 fluorometer was used for DNA quantification in ng/µl. *r* value range between 0.93 – 0.97.

3.4.4.1.1.2 Isolation of DNA from AK samples for WES

DNA was extracted from dysplastic epidermis and adjacent normal skin from the selected FFPE AK samples (N = 5). Between 40 and 55 sections (10 µm thick) were required per sample, depending on the cellularity and the size of the target area. A QIAamp DNA FFPE Tissue Kit was used for DNA extraction according to manual's protocol with minor modifications (section 2.2.9), including increasing the volume and incubation time of the sections with proteinase K (i.e. up to 80 µl of proteinase K) to ensure the complete lysis of the sections. All samples gave a total DNA concentration of ≥ 500ng / 60 µl as presented in table 3.5, which based on previous experience within the department, was likely to be sufficient for WES and to give reliable results, despite the recommended DNA concentration of 1-3 µg in the Agilent SureSelect V5 library preparation kit.

Table 3.5 Concentration of DNA isolated from FFPE AK samples (N=5) for WES study. DNA concentration was measured using a Qubit 2.0 fluorometer. Although the recommended DNA quantity for WES using the Agilent SureSelect V5 library preparation kit is 1-3 µg, samples with a total DNA concentration (in 60µl) ≥ 500ng were considered likely to be suitable for next generation sequencing.

Sample ID	Dysplastic tissue area		Matched adjacent normal looking skin area	
	DNA concentration in ng/µl	DNA concentration in ng/60 µl	DNA concentration in ng/µl	DNA concentration in ng/60µl
28	15.2	912	14	840
30	42.6	2556	17	1020
51	97	5820	57.7	3462
54	27.3	1638	26.6	1596
68	35.4	2124	14.54	872.4

3.4.4.1.2 DNA quality assessment

To evaluate DNA quality, the purity and degradation of the extracted DNA was assessed. A Nanodrop spectrophotometer was used to assess DNA purity according to section 2.2.10.1. This technique relies on the fact that DNA molecules absorb light at 260 nm wavelength and any contamination (organic or nonorganic) will cause a deviation in spectrum of light absorbance, thus the 260/280 and 260/230 absorbance ratios indicate the purity of the DNA. Using this technique, the DNA extracted from the AK samples (N = 5) was considered to be of appropriate purity for subsequent genetic analysis and the result are presented in figure 3.21.

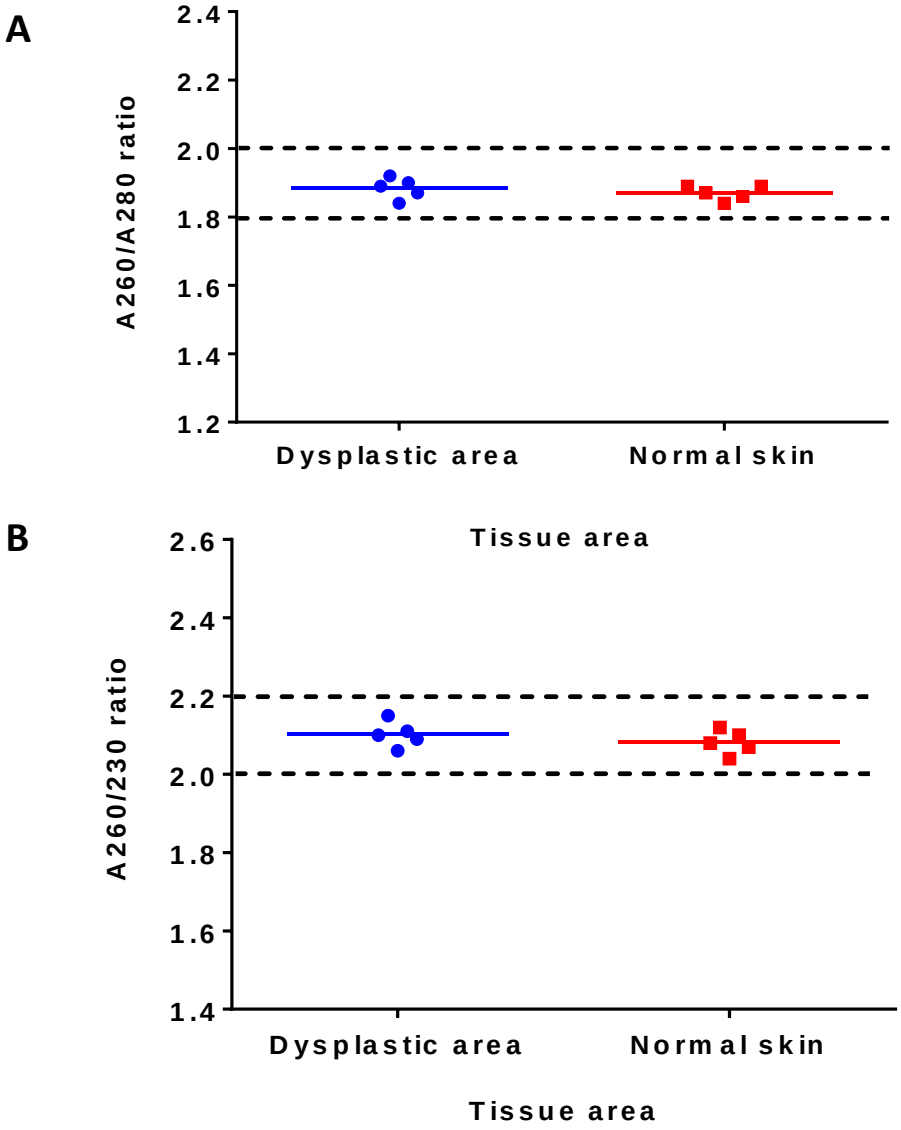


Figure 3.21 DNA quality assessments for the DNA extracted from AKs (N=5). UV absorbance spectra of DNA using Nanodrop spectrophotometer from the dysplastic and normal looking skin from the AK tissue sections are shown in this figure. **A**: DNA A260/A280 ratio; values between 1.8 and 2.0 indicate high purity DNA with minimal protein contamination. **B**: DNA A260/A230 ratio; values between 2.0 and 2.2 indicate high purity DNA with minimal solvent contamination.

3.4.4.1.2.1 Agarose gel electrophoresis (protocol optimisation step)

To assess DNA fragmentation, PCR and agarose gel electrophoresis were performed on the DNA extracted from the AK samples (sections 2.2.11.1 & 2). Good quality DNA usually appears as a high molecular weight band on the agarose gel (Xuan et al., 2013), while fragmented DNA appears as a smear on the gel (Hostetter et al., 2010). As the DNA in the study was obtained from FFPE sections, the presence of smears rather than a single high

molecular weight band on the agarose gel was expected. For comparison, blood DNA was used as a control.

Moreover, as obtaining sufficient DNA concentration from AKs was not easy and required multiple 10 µm tissue sections, in order to establish the protocol for DNA extraction from FFPE tissue, breast skin was utilised. Samples were loaded and run on 2% agarose gels according to the protocol in section 2.2.11.2. Information on the DNA samples from breast skin (N = 5) is summarised in table 3.4 (above) and from blood (N = 5) in table 3.6. The agarose gel results showed a smear from the FFPE DNA samples, however, faint bands of high molecular weight DNA were also observed. The result is shown in figure 3.22.

Table 3.6 Concentration and purity results of DNA extracted from blood samples (N =5). The measurements were obtained from the Qubit 2.0 fluorometer (for concentration) and Nanodrop spectrophotometer (for purity as assessed by A260/A280 and A230/A260 ratios). DNA was diluted 1:10, thus a comparable concentration of DNA to the DNA concentration obtained from breast skin samples was used.

Sample ID	DNA concentration in ng/µl (Qubit)	DNA concentration in ng/µl after dilution 1 in 10 (Qubit)	A260/A280 ratio	A260/A230 ratio
1	183	18.3	1.88	2.09
2	175	17.5	1.82	2.04
3	182	18.2	1.94	2.18
4	157	15.7	1.97	2.15
5	197	19.7	1.87	2.05

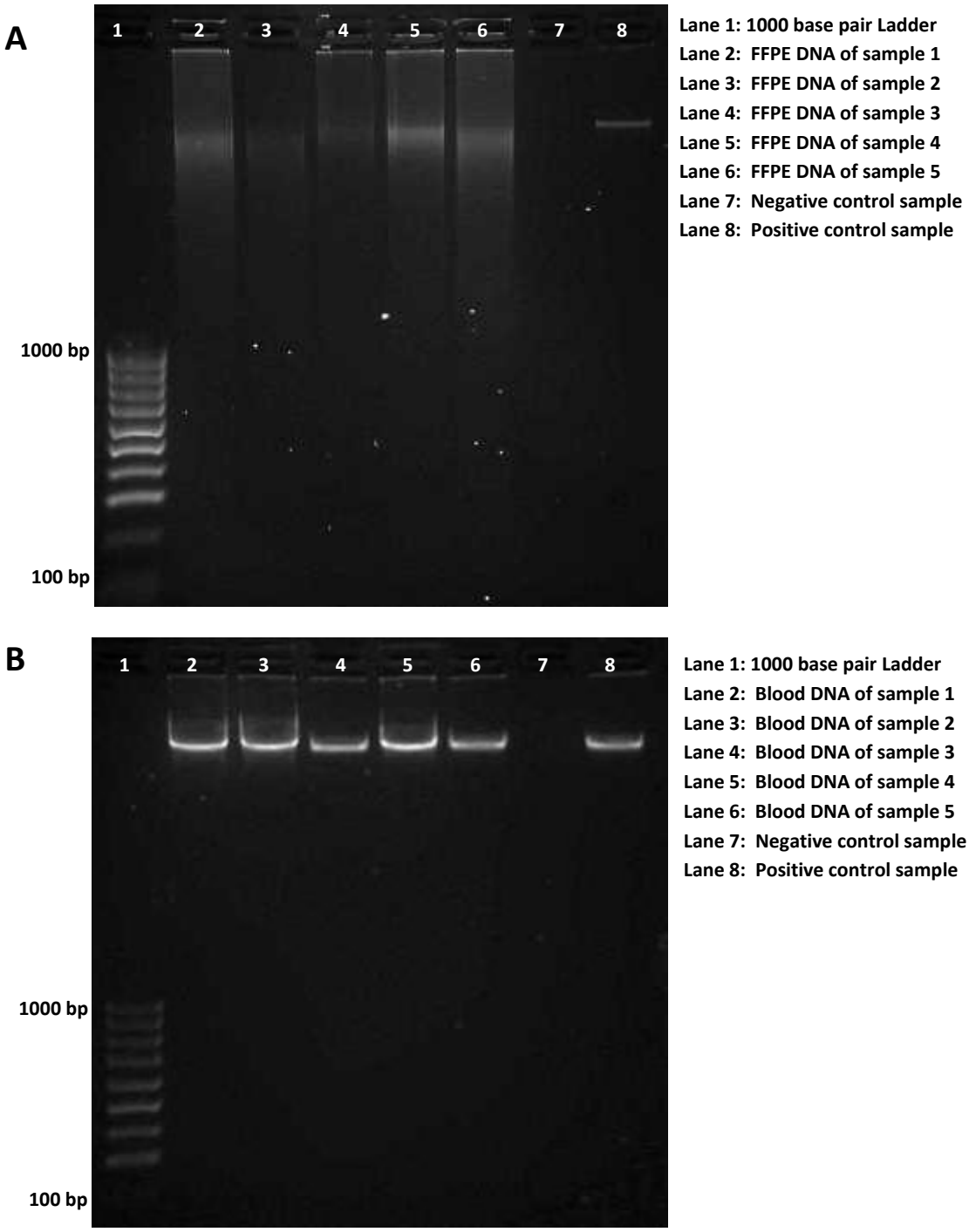


Figure 3.22 Gel electrophoresis of genomic DNA. **A:** DNA samples extracted from FFPE breast skin (as detailed in table 3.4). **B:** Blood DNA samples (as detailed in table 3.6). On both gels, lane 1 contains DNA ladder (bands 100, 200, 300, 400, 500, 600, 700, 800, 900 and 1,000 bp). DNA samples (from FFPE in A and blood in B) were loaded onto lanes 2 to 6. Lanes 7 and 8 represent negative and positive controls respectively.

3.4.4.1.2.2 PCR DNA quality assessment

3.4.4.1.2.2.1 *MC1R* gene *PCR*

PCR was used to determine whether the DNA extracted from FFPE samples (and from blood as control) could be amplified as this step is necessary during WES. The *MC1R* gene was used as an initial template for the PCR and was conducted on the FFPE breast and blood DNA samples that were used in section 3.4.4.1.2.1 according to the methods detailed in section 2.2.11.1. Blood DNA samples gave the 954 bp target size band, but FFPE DNA failed to give any results (the experiment was repeated 3 times to exclude any reaction errors). This result presented in figure 3.23 and indicated that much of the DNA isolated from FFPE tissue was likely to be fragmented.

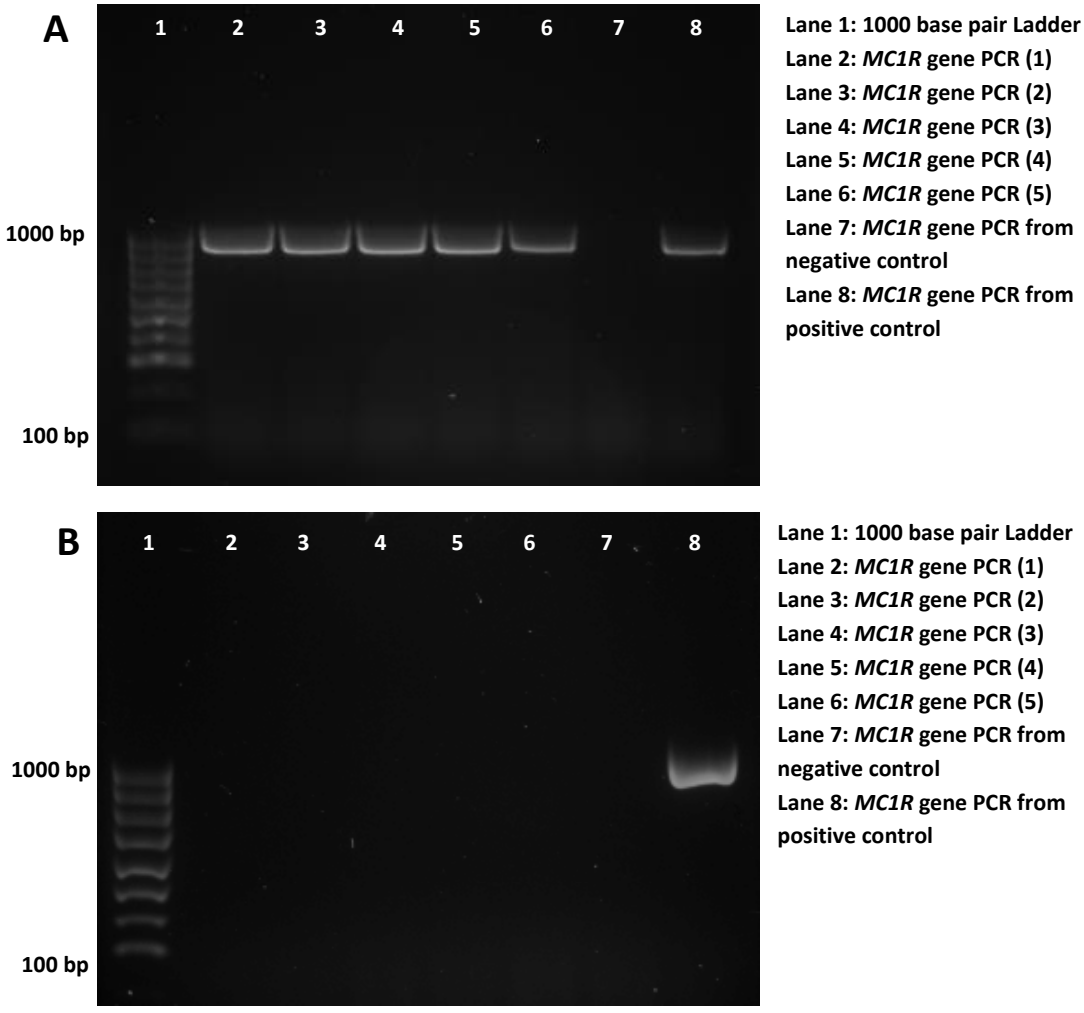


Figure 3.23 Gel electrophoresis for *MC1R* gene PCR products. PCR was performed using *MC1R* gene forward and reverse primer pairs. **A:** Blood DNA samples (as detailed in table 3.6) and **B:** DNA samples extracted from FFPE breast skin (as detailed in table 3.4). The expected PCR product of *MC1R* gene is 954 base pairs in length. Lane 1 on the agarose gels contains DNA ladder (bands 100, 200, 300, 400, 500, 600, 700, 800, 900 and 1,000 bp), whereas lanes 2 – 6 contain DNA samples from blood in A and FFPE skin in B. Lanes 7 and 8 represent negative and positive controls respectively. bp = base pair.

3.4.4.1.2.2.2 Panel (Multiplex) PCR for FFPE breast skin samples

Panel PCR is a multiplex based PCR that uses different primer pairs that amplify different lengths of DNA. In this project, the expected lengths of the panel PCR products were 100, 200, 300, 400 and 600 bps. According to the manufacture's protocol for the QIAamp DNA FFPE Tissue Kit, DNA up to 650 bps from FFPE sections can be obtained. The PCR conditions were as listed in section 2.2.11.1. The panel PCR was first conducted on FFPE breast DNA samples that were used in section 3.4.4.1.2.1 and the products electrophoresed on an agarose gel. From the results in figure 3.24, it can be seen that PCR amplicons of up to 600 bps in size can be amplified from FFPE DNA samples.

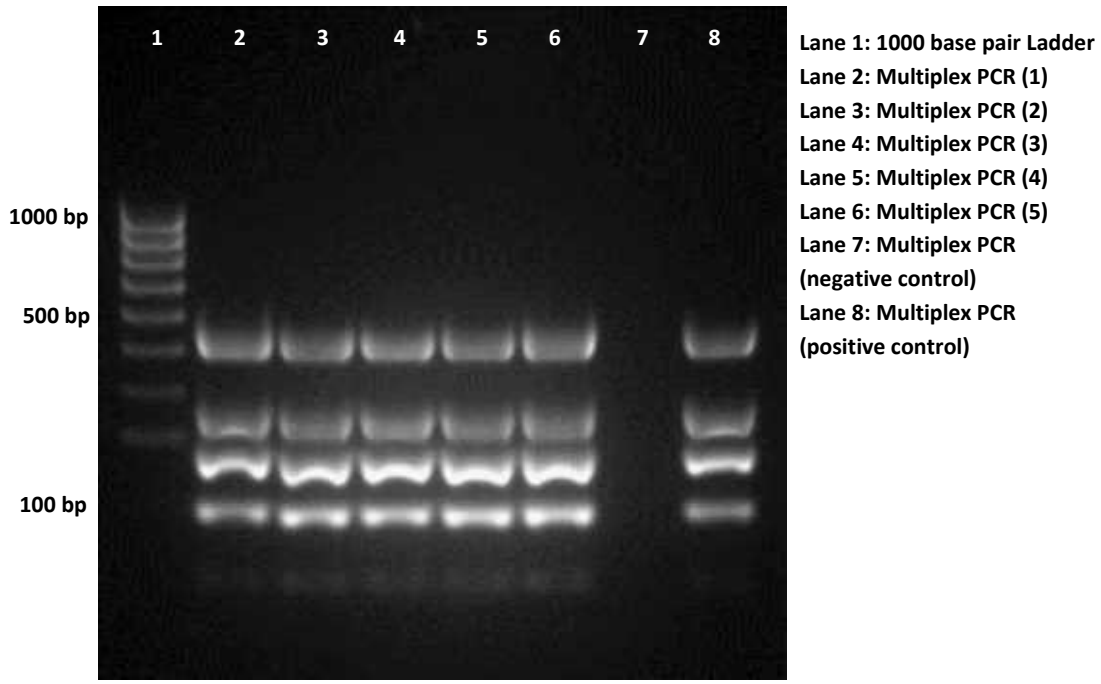


Figure 3.24 Gel electrophoresis for panel (multiplex) PCR products. PCR was performed on DNA samples extracted from FFPE breast skin (as detailed in table 3.4) using multiplexed primers that amplify DNA to give PCR products of different lengths (100, 200, 300, 400 and 600 base pairs). Lane 1; DNA ladder (100, 200, 300, 400, 500, 600, 700, 800, 900 and 1,000 bp). FFPE skin samples are labelled as 1 – 5 in lanes 2 – 6. Negative control (no DNA in PCR, lane 7) and positive control (blood DNA, lane 8).

3.4.4.1.2.2.3 Panel (Multiplex) PCR for FFPE AK samples

Panel PCR was performed on the DNA isolated from the dysplastic and matched adjacent normal skin areas from FFPE AK samples (N = 5). In contrast to FFPE DNA samples from breast skin presented in figure 3.24, the panel PCR results of AK DNA showed variation in

the maximum length of the DNA fragments in bps. While the largest PCR product for samples 28 and 30 was 300 bp, the maximum PCR product for samples 51 and 54 was 400 bp and was 600 bp for sample 68. The exact reason for these differences is not clear. FFPE DNA from breast skin was used soon after being formalin fixed paraffin embedded while the AK blocks had been stored following embedding for a variable length of time up to several years (the oldest blocks were paraffin embedded in 2003). Although panel PCR is not accurately quantitative, the brightness of the PCR bands on the agarose gel can indicate which bands contain more PCR products. Thus, from the gel results it can be estimated that the majority of DNA products within the samples were likely to be 300 bps. This suggested that the DNA isolated from the FFPE AKs was likely to be suitable for WES study because the first step in WES involves library prepartion which uses DNA of 150-200 bps length. These results are summarised in figure 3.25.

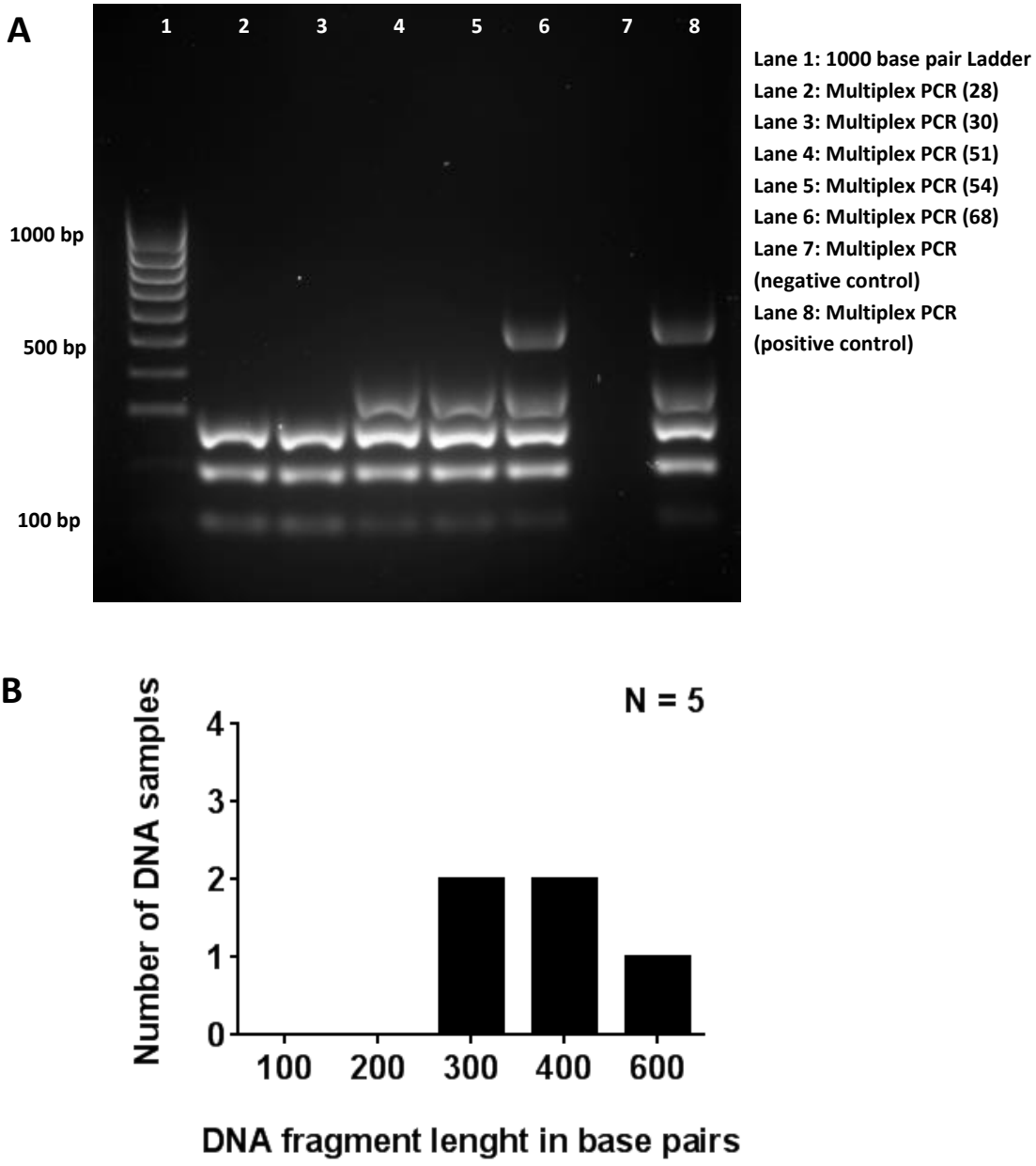


Figure 3.25 Panel PCR products of the selected AKs sections for WES (N=5). **A:** Agarose gel electrophoresis results. Panel PCR product from DNA samples extracted from lesional (dysplastic) area of AKs using a variety of primers that amplify DNA to give PCR products of different lengths (100, 200, 300, 400 and 600 base pairs). Lane 1; DNA ladder. Lanes 2 – 6; FFPE samples (labelled according to their sample identification number (28, 30, 51, 54 & 68). Lane 7; negative control (no DNA in PCR. Lane 8; positive control (blood DNA in PCR).**B:** Bar graph representation for the multiplex PCR product from figure A. bp = base pair.

3.4.5 Whole exome sequencing (WES)

3.4.5.1 Selection of library preparation kit and sequencing system

There are a number of whole exome capture kits for human exome sequencing supplied from different providers (e.g. Illumina, Agilent, and BGI (Beijing Genomics Institute)) (Meldrum et al., 2011, Xuan et al., 2013). In order to determine the most appropriate exome capture kit for this project, a list of genes that were known to be altered in AKs, cSCC, Bowen's disease and BCC was created based on information from (Campbell et al., 1993b, Durinck et al., 2011, Griewank et al., 2013, Jayaraman et al., 2014, Kanellou et al., 2008, Lee et al., 2000, Mauerer et al., 2011, Park et al., 1996, Ping et al., 2001, Saridaki et al., 2003, Sottoriva et al., 2015, Soufir et al., 1999, South et al., 2014, Taguchi et al., 1998, Takata et al., 1997, Ziegler et al., 1994) (appendix tables 8.10 and 8.11). As different exome enrichment kits vary in the percentage of coverage of different genes, it was important to know which kit would give the best coverage of most, if not all, the genes of interest. Two panels of genes were investigated (appendix table 8.12 A and B); the first panel is the most relevant as it represents genes that have been shown to be altered in AKs or other related lesions (cSCCs and Bowen's disease). The second panel comprises genes that have been shown to be mutated in BCCs.

The process of examination of the best gene/kit match was done in collaboration with Sarah Ennis (Professor of Genomics, Genetic Epidemiology and Genomic Informatics Group, Human Development and Health, Faculty of Medicine, University of Southampton). In addition to the gene coverage, other factors including the availability and the cost of the kit were also taken in consideration. The result of this matching is summarised in (appendix table 8.12 A and B). Although the BGI kit had the lowest prices (£500/ sample at the time of assessment), the kit was omitted as it also had the lowest percentage of gene coverage. The percentage of gene coverage of SureSelect v5 (Agilent v5) and Illumina TruSeq kits were comparable as well as the cost (£700 and £650 respectively). However, after various considerations including coverage, cost, availability and the recommendation of Prof Sarah Ennis, the SureSelect v5 (Agilent v5) exome capture kit was selected. Using SureSelect v5 (Agilent v5) capture kit and Illumina HiSeq 2000 sequencing system, WES was performed in the Wellcome Trust Centre for Human Genetics, Oxford.

3.4.5.2 Quality assessment and somatic variant calls

Sequencing results were received in the form of fastq files. The preliminary analysis for read mapping and variant calling was undertaken by Dr Reuben Pengelly, Genetic Epidemiology and Genomic Informatics Group, Human Development and Health, Faculty of Medicine, University of Southampton. For sequencing data quality assessment, Novoalign v2.08.02 and ANNOVAR v2012Jun21 pipelines were used (following the standard University of Southampton Genetic Epidemiology Group pipeline settings). The parameters used for the quality assessment performed by Dr Reuben Pengelly included candidate gene coverage and genotype quality control.

A. Candidate gene coverage

The depth of coverage of the samples (N = 10, i.e. 5 dysplastic and 5 matched adjacent normal looking skin) was estimated using BEDTool v2.13.2 (table 3.7 – part 1). The mean depth of coverage at 20X was 90.9%. The sequencing coverage for DNA samples from dysplastic epidermis (labelled A) was more than that from matched adjacent normal looking skin (labelled B). This was because of the experimental design; to increase the depth of coverage for the lesional AK samples, the 5 DNA samples of dysplastic area were loaded onto one lane of the sequencing plate (the maximum capacity of a lane is 8 samples), but to minimise costs the DNA samples from matched adjacent normal looking skin were loaded along with three samples from an unrelated project onto another sequencing lane.

B. Genotype quality control

As an important quality assessment step, samples matching were conducted. For every individual, approximately, 24,000 - 25,000 variants are expected on WES, therefore a matrix of genotype concordance across all samples was conducted. Unrelated samples are expected to share ~44% of variants, whereas this should be > 95 % for samples from the same individual in most cases (table 3.7 – part 2).

Table 3.7 WES quality assessment and somatic variant calls. (**Part 1**) Depth of coverage for WES results (N=10) of lesional (labelled A) and normal (labelled B) DNA from each sample. The base pairs within the target reads for all samples showed more than 80% depth of coverage at 20x depth of coverage. (**Part 2**) Sample matching for WES results using Single Nucleotide Polymorphism (SNP) analysis comparing lesion (labelled A) and adjacent normal looking skin (labelled B) of AK samples. Samples from the same individual are shown in green, with variation between normal and lesional samples > 95% in most of the cases. Data analysis was undertaken by Dr Reuben Pengelly.

Part 1

Sample ID	Total reads	Target reads	Target At 20x (%)	Target at 10x (%)	Target at 5x (%)	Target at 1x (%)
AK28A	71165395	65839630	93.65	97.82	98.91	99.47
AK28B	49358924	45478789	88.96	96.68	98.57	99.44
AK30A	61318469	55999319	92.28	97.38	98.75	99.44
AK30B	40657578	36928395	84.43	95.43	98.21	99.39
AK51A	68045252	61731792	92.78	97.45	98.69	99.35
AK51B	59378977	53945124	92.5	97.52	98.75	99.4
AK54A	76425854	70353140	94.42	97.77	98.8	99.45
AK54B	49520774	45596017	89.63	96.85	98.63	99.44
AK68A	83122511	76353039	95.8	98.32	99	99.41
AK68B	40501403	36964680	84.32	95.5	98.21	99.32

Part 2

	AK28A	AK28B	AK30A	AK30B	AK51A	AK51B	AK54A	AK54B	AK68A	AK68B
AK28A	100	98.31	41.18	40.54	39.92	40.65	41.08	40.96	39.08	40.83
AK28B	96.35	100	44.22	43.91	43.08	44.05	44.34	44.4	42.11	44.19
AK30A	42.13	41.9	100	97.17	40.57	41.32	41.8	41.53	40.07	44.79
AK30B	44.8	44.94	96.32	100	43.25	44.2	44.34	44.3	42.53	44.73
AK51A	40.74	40.72	40.47	39.94	100	97.39	40.1	39.88	39.69	41.86
AK51B	44.53	44.68	44.24	43.8	98.79	100	44.15	43.98	43.66	45.98
AK54A	43.83	43.81	43.58	42.8	41.92	43	100	98.05	40.82	42.66
AK54B	44.55	44.73	44.15	43.6	42.5	43.67	96.91	100	41.53	43.6
AK68A	44.38	44.29	44.48	43.7	44.17	45.27	43.45	43.36	100	96.11
AK68B	44.31	44.41	44.33	43.93	44.51	45.55	43.4	43.5	97.02	100

3.4.5.3 Overview on exome sequencing results

Using the Agilent SureSelect V5 library preparation kit, 21,522 genes were targeted in the WES of AK samples (N = 5) along with matched adjacent normal looking skin. The total number of exonic somatic alterations (silent and non-silent) was 7,270 and the median number of mutations per AK was 1,723 (range 267 – 2,492). The median number of mutated genes per AK was 1,275 (range 194 – 1,688). A total of 5,012 non-silent somatic alterations (median of 1,197 per AK, range 203 – 1,699) were seen, corresponding to a median total somatic mutation burden of 34.5 (range 5.3 – 49.8) mutations /megabase of DNA. UV signature mutations (C > T and CC >TT) transition base substitutions were present in all AKs, with a median C > T mutation frequency of 80.2 % (range 51.4 – 80 %)

and a median CC > TT mutation frequency equal to 4 % (range 1 – 6%). Missense mutations were also common with a median frequency of 53.7 %. Overall, the ratio of non-silent to silent (synonymous) events was around 4:1 in the dysplastic and matched adjacent normal looking skin respectively. The results are summarised in figure 3.26.

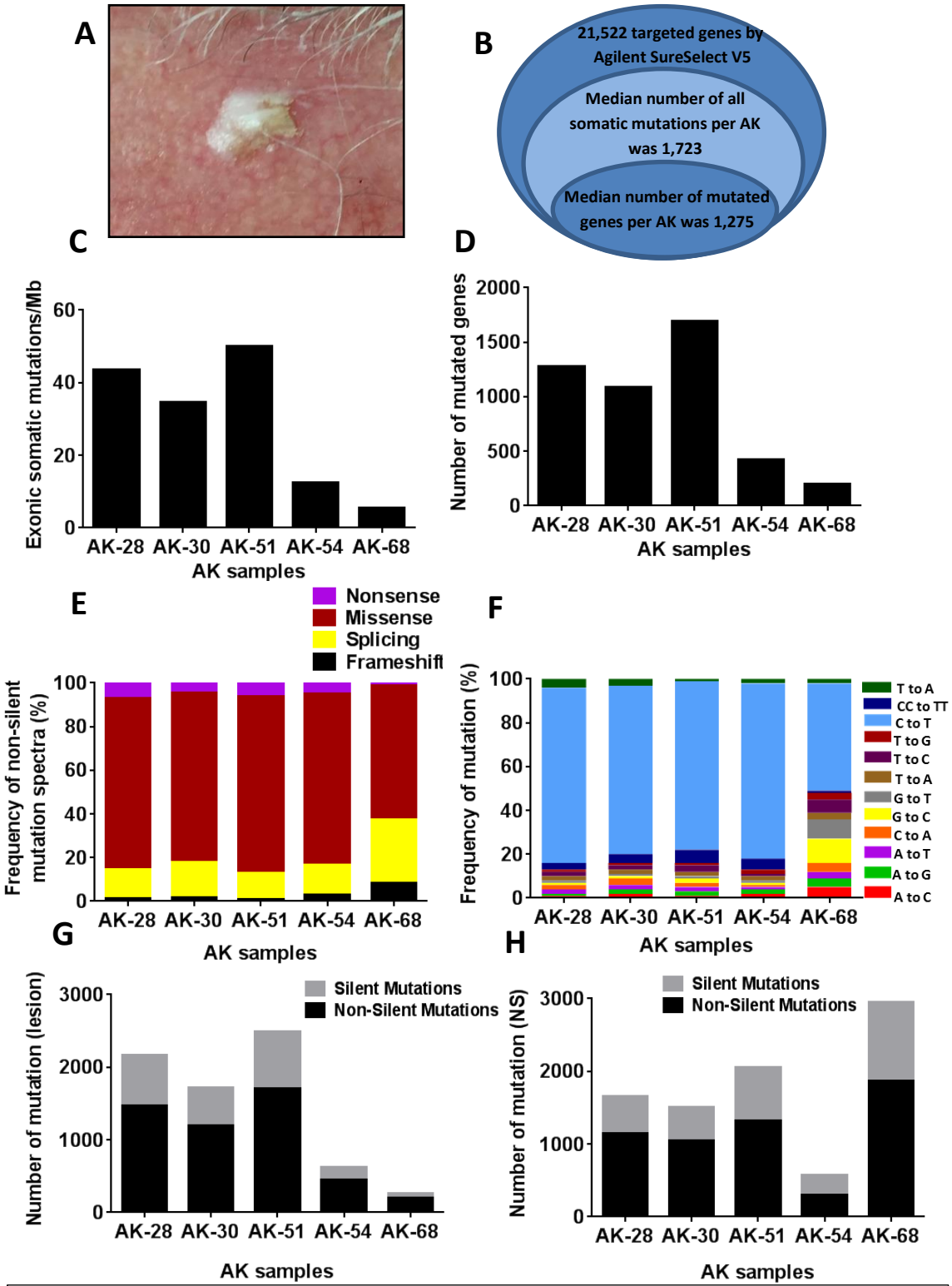


Figure 3.26 Integrated overview on the mutation rate and type within AKs. **A:** Clinical picture of AK lesion. **B:** Total number of genes captured and sequenced by Agilent SureSelect V5 and Illumina HiSeq 2000 platform respectively. The median number of mutated genes (1,275) and median number of all somatic mutations (1,723) per AK are indicated. **C**: Mutation burden per AK; the median somatic mutation rate was 34.5 mutations/megabase. **D**: Number of mutated genes per AK (range 194 – 1,688). **E&F**: The frequency of mutation spectra within WES data showing missense mutations and C >T alterations as the most common changes. **G&H**: The frequency of silent (synonymous) and non-silent genetic alterations within the five AKs and matched normal-looking skin respectively. The ratio of the median non-silent: silent genetic alterations are 4:1 in the dysplastic area and in the matched normal skin. AK = actinic keratosis. NS = normal skin. Mb = megabase.

3.4.5.4 Characterisation of mutated genes according to the frequency of mutation within WES results

The AK WES results showed a high mutation burden with high number of mutated genes, therefore identification of candidate driver genes was challenging. Previous work by South et al. (2014) utilised an analytical approach that involved ordering detected mutated genes according to the frequency of mutation of those genes within the study. This analytical approach from South et al. (2014) study was adopted for analysis of the WES AK data. Thus, mutated genes were ordered according to their frequency *TP53* was ranked highest (mutated within all 5 tested AKs). Other genes demonstrated to be mutated in high frequency within WES AK data include *TTN, CCDC168, MUC16, PCLO, and UNC79* (within 5, 4, 4, 4, and 4 of the AKs respectively). False positively mutated genes were picked up using the Network of Cancer Genes webpage (NCG5.0) (http://ncg.kcl.ac.uk/) (Genes were classified as false positively mutated based on calculations by Lawrence et al. (2013)). In the WES AK study (N = 5) , out of 94 genes that were mutated in ≥ 40% of the AK WES data (i.e. 2 or more AKs), of which 11 genes including *TTN, CCDC168, MUC16, PCLO, and UNC79* were reported as false positive mutated genes based on their size from a meta-analysis of whole exome and genome sequencing data (the large size of the genes make them highly vulnerable to be mutated) (Lawrence et al., 2013). These genes were thus discounted in downstream analysis.

3.4.5.5 AK shares most of the highly mutated genes with cSCC

Previous exome sequencing studies on cSCCs identified certain genes which are highly mutated within those lesions (Li et al., 2015, Pickering et al., 2014, South et al., 2014). Thus, in the current study, the previously reported frequently mutated genes within cSCC were compared with the WES results from AKs. In addition, the Catalogue of Somatic Mutations in Cancer (COSMIC) (http://cancer.sanger.ac.uk/cosmic) was also used to check if any of the highly mutated genes in AKs had been previously reported in skin cancer and/or other human cancers. The results are summarised in appendix table 8.13. Out of a median of 1,275 mutated genes, 94 genes were mutated in ≥ 40% of the AK WES data (i.e. 2 or more AKs), of which 43 genes were mutated within the previous exome sequencing cSCC studies (Li et al., 2015, Pickering et al., 2014, South et al., 2014). An additional 36 genes were mutated in 20% of the AKs (i.e. only in 1 AK) and 11 genes with no mutation in the WES results of the AKs were reported to be mutated in cSCC.

Moreover, 84 out of the 94 genes were reported previously in skin cancer and/or other human cancers in COSMIC cancer gene census.

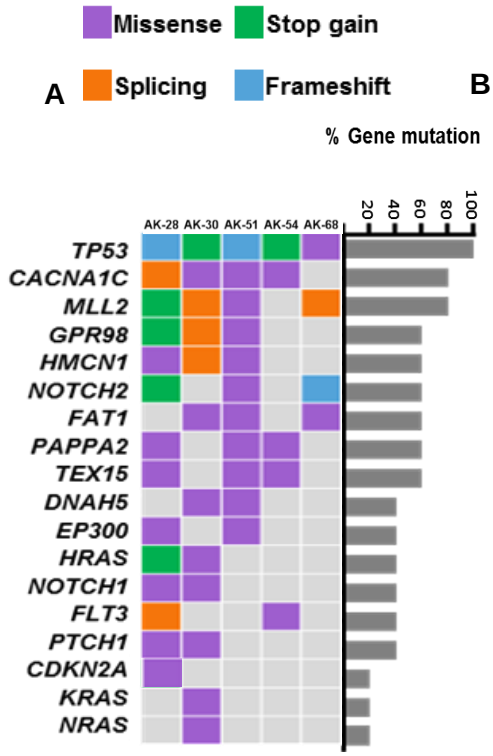
3.4.5.6 Selection of genes for target enriched sequencing (variant validation)

Selected genes were prioritized according to a set of filtration criteria (Lee et al., 2014, South et al., 2014). At first, WES data were filtered, thus, only genes that have non-silent somatic mutations that shown to be damaging (according to SIFT and PolyPhen2 algorithms) were considered as a potential target for target enriched resequencing (synonymous alteration and previously reported variants in the dbSNP Build 137 was excluded from the analysis). Then, a gene list with all genes that were mutated within the study (N = 5) was generated. The list was re-rearranged according to the frequency of mutated genes within the AKs and assessed if any of these genes were reported previously in WES data from cSCC (Li et al., 2015, Pickering et al., 2014, South et al., 2014) and in Catalogue of Somatic Mutations in Cancer (COSMIC) (Forbes et al., 2010) database as stated previously.

As the plan was to analyse the mutation frequency within highly mutated genes on a larger sample size of precancerous skin conditions (chapters 4 and 5), 18 genes out of the 141 mutated genes in appendix table 8.13 were selected for target enriched sequencing. It was decided that for a gene to be selected, it should (i) be mutated in ≥ 40% of AK WES data, (ii) not have been reported as false positively mutated according to the network of cancer gene webpage, (iii) have been reported as mutated previously in at least one of the cSCC exome sequencing studies, and/or (iv) reported previously in the COSMIC cancer gene census as a cancer gene. Although most of the 18 genes selected were mutated in $\geqslant$ 40% AK samples, some of the selected genes appeared in only 20% of AKs (*KRAS, NRAS* and *CDKN2A*) as these latter genes have been demonstrated previously to be mutated in cSCC, in AK, and also in COSMIC cancer gene census. Although the *PTCH1* gene is mainly considered relevant to the development of BCC rather than cSCC, it met the selection criteria above and a recent study by Criscione et al. (2009) showed that BCCs can develop at sites of previous AK lesions; therefore *PTCH1* was included in the gene list. The selected genes appear highlighted in green in appendix table 8.13.

3.4.5.7 Landscape of mutation spectrum and mutated gene frequency of the 18 selected genes within the AK WES results

TP53 mutations were detected in 5 AKs (100%), making *TP53* gene the most common mutated gene within this study. This confirms a previous report of *TP53* point mutations in AKs by Nelson et al. (1994), Nindl et al. (2007), Park et al. (1996), Soufir et al. (1999), Taguchi et al. (1998), Ziegler et al. (1994). Interestingly, *MLL2* and *CACNA1C* genes were mutated in 4 of the 5 AKs. *GPR98, HMCN1, NOTCH2, FAT1, PAPPA2* and *TEX15* were mutated in 3 out of 5 (60%) AK samples and *DNAH, EP300, HRAS*, *NOTCH1*, *FLT3* and *PTCH1* were mutated at 40% frequency. In keeping with previous studies on cSCC (figure 3.27), *CDKN2A*, *NRAS* and *KRAS* were mutated at a low frequency of 20% within the AKs. The results are presented in figure 3.27.



AK (current study)	South et al., 2014	Durinck et al., 2011	Lee et al., 2014	Pickering et al., 2014	Li et al., 2015	Chitsazzadeh et al., 2016	Ping et al., 2001
100	64	87.5	42	59	79	57	
80	≥ 40		1	24V/39			
80	≥ 40			69.2	17.2	14	
60	≥ 40		3	71V/39			
60	≥ 40	37.5¥	2	37V/39			
60	63	50	34	54	69¶	28.5	
60	≥ 40		41	43.6		28.5	
60	≥ 40	50¥		34V/39			
60	≥ 40	25¥		20V/39			
40	≥ 40	50¥		68V/39			
40	25	25		11V/39	34.5		
40	16	12.5	6	20.5	3		
40	75	75	54	59	69¶	14	
40		37.5¥		10V/39	2V/29		
40	25	12.5¥	15	6V/39	6V/29		25
20	23	25	28	43.6	48		
20	13		3	2V/39	3		
20	5	12.5¥		5.1			

Figure 3.27 WES AK data on 18 genes selected for target enriched sequencing (in chapters 4 and 5). **A:** Heatmap representation of the number and types of mutations within these genes in each AK. The corresponding bar graph on the right represents the mutation frequency of each of these genes within the group of AKs presented as percentage of the total number of samples. **B:** The table on the right highlights the frequency of somatic mutations within the same gene in the heatmap in cSCC using WES (Chitsazzadeh et al., 2016, Durinck et al., 2011, Pickering et al, 2014, South et al., 2014), Targeted enriched sequencing (TES) (Lee et al, 2014, Li et al., 2015,) and Sanger Sequencing of cSCC (Ping et al., 2001). ¥ Sign indicates frequency calculated from the supplementary data of relevant study. ¶ Sign indicates data combined from the study publication and corresponding supplementary data. V/39 & V/29 represent the number of variants detected within the reported study over the total number of samples (the frequency of somatic mutation could not be calculated as the number of samples with each of the variants was not provided in the relevant study data). WES = whole exome sequencing. TES = target enriched sequencing.

Following completion of the WES analysis of the 5 AKs presented in this thesis, Chitsazzadeh et al. (2016) published the results of a study identifying driver genes within cSCC, AK and normal skin across-species (human and mice). While their results were not considered in the process of selection for the 18 genes, the WES results on the tested AKs by Chitsazzadeh et al. (2016) were compared to the WES AK data from the current project and the results of this comparison will be discussed in later in section 3.5.

Information on the position, the type of the mutations and the amino acid changes on each of the resultant protein from these 18 genes is summarised in figure 3.28 (A-E) and in appendix table 8.14. The somatic mutations were scattered over the protein domains with some proteins showing mutation predominantly in certain domains (examples

include p53, NOTCH1, CACNA1C and GPR98). Unlike melanoma where hotspot mutations in *BRAF* and *NRAS* have been reported (Hodis et al., 2012), no hotspot mutations were observed in any of the 18 genes in the group of AKs, however, this may be due the small size of the AK study.

Note that, as figure 3.28 extends over 5 pages (parts A – E), a legend has been added to each part of the figure for ease of reading.

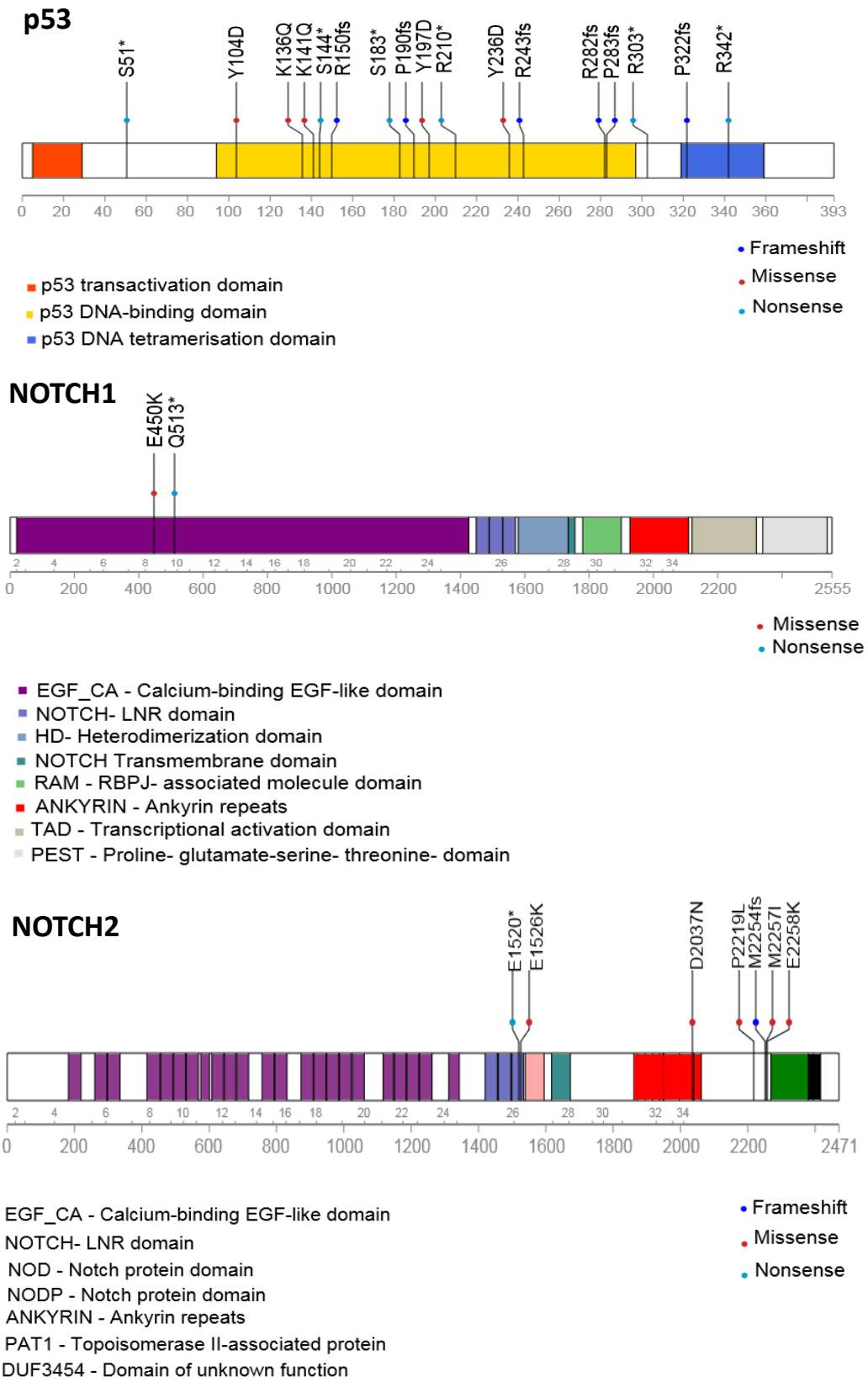


Figure 3.28 (A) Information on the amino acid changes resulting from the mutations in genes selected for target enriched sequencing. The type of somatic mutations (missense, nonsense, frameshift and deletion) in the WES AK data is also shown. Protein names are as per (NCBI)(www.ncbi.nlm.nih.gov/protein/). The diagram was generated using Pediatric Cancer Genome Project (www.explore.pediatriccancergenomeproject.org/proteinPainter).

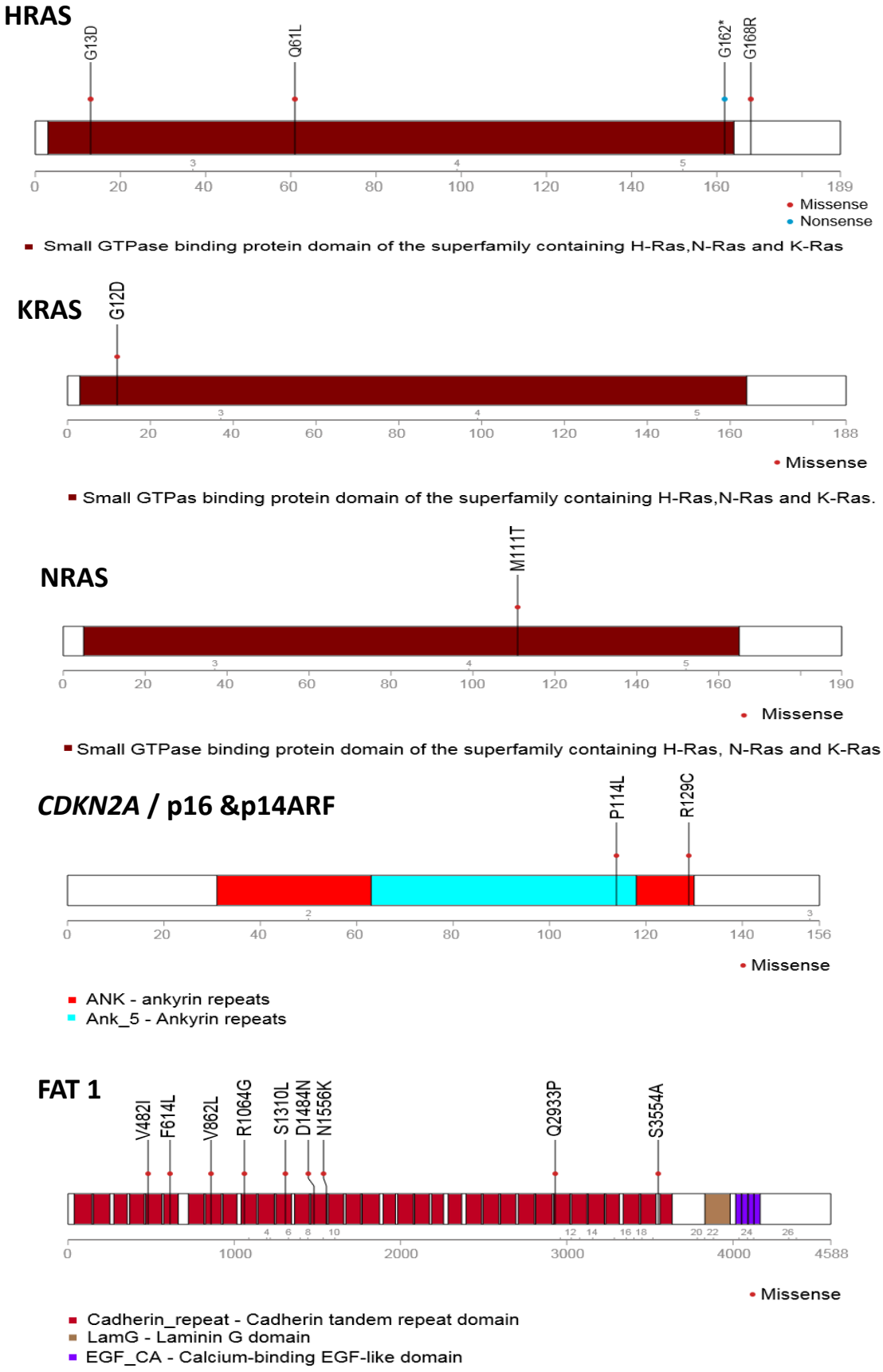
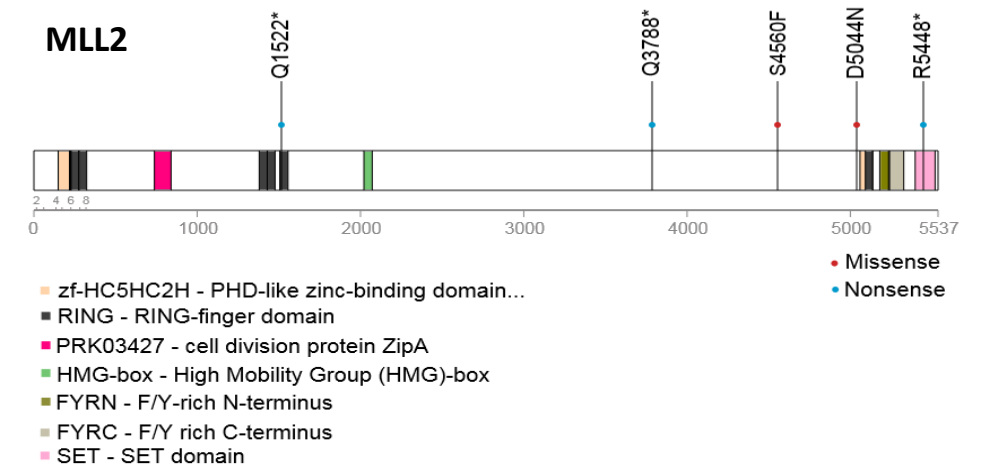
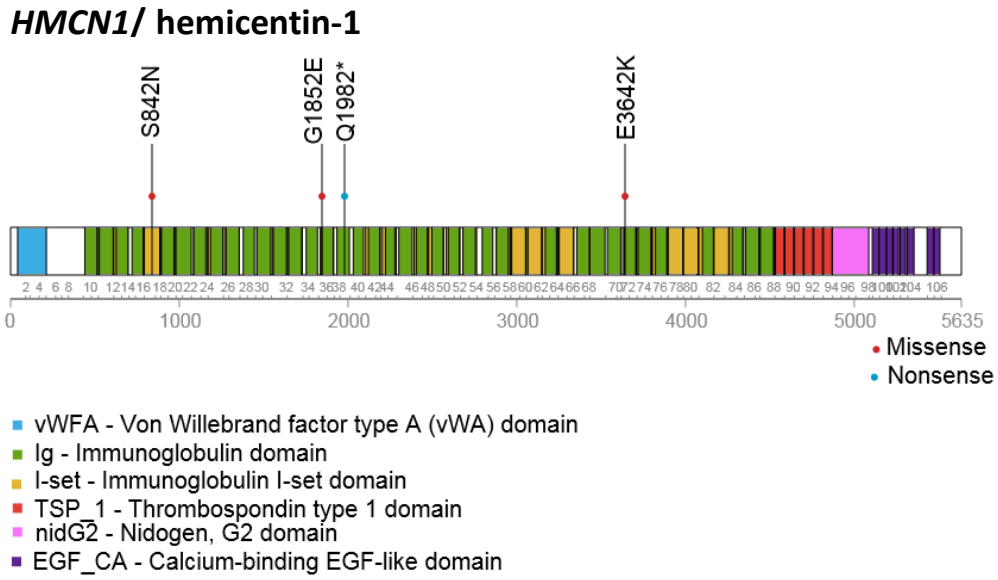


Figure 3.28 (B) Information on the amino acid changes resulting from the mutations in genes selected for target enriched sequencing. The type of somatic mutations (missense, nonsense, frameshift and deletion) in the WES AK data is also shown. Protein names are as per (NCBI) (www.ncbi.nlm.nih.gov/protein/). The diagram was generated using Pediatric Cancer Genome Project (www.explore.pediatriccancergenomeproject.org/proteinPainter).





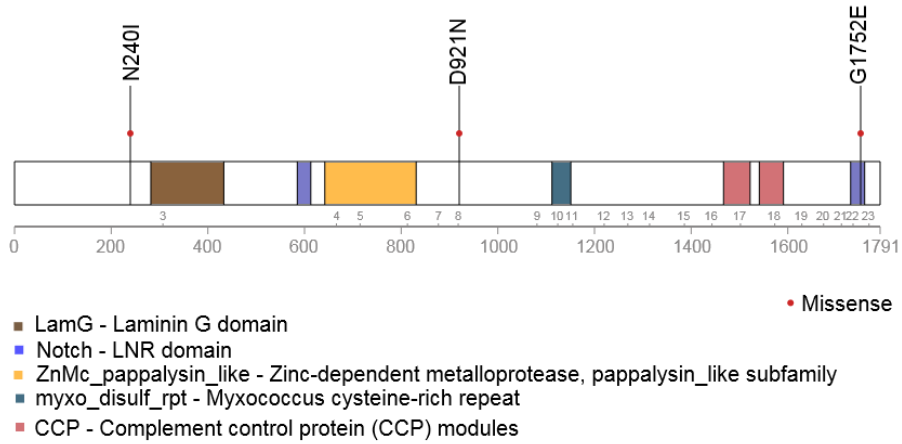


Figure 3.28 (C) Information on the amino acid changes resulting from the mutations in genes selected for target enriched sequencing. The type of somatic mutations (missense, nonsense, frameshift and deletion) in the WES AK data is also shown. Protein names are as per (NCBI) (www.ncbi.nlm.nih.gov/protein/). The diagram was generated using Pediatric Cancer Genome Project (www.explore.pediatriccancergenomeproject.org/proteinPainter).

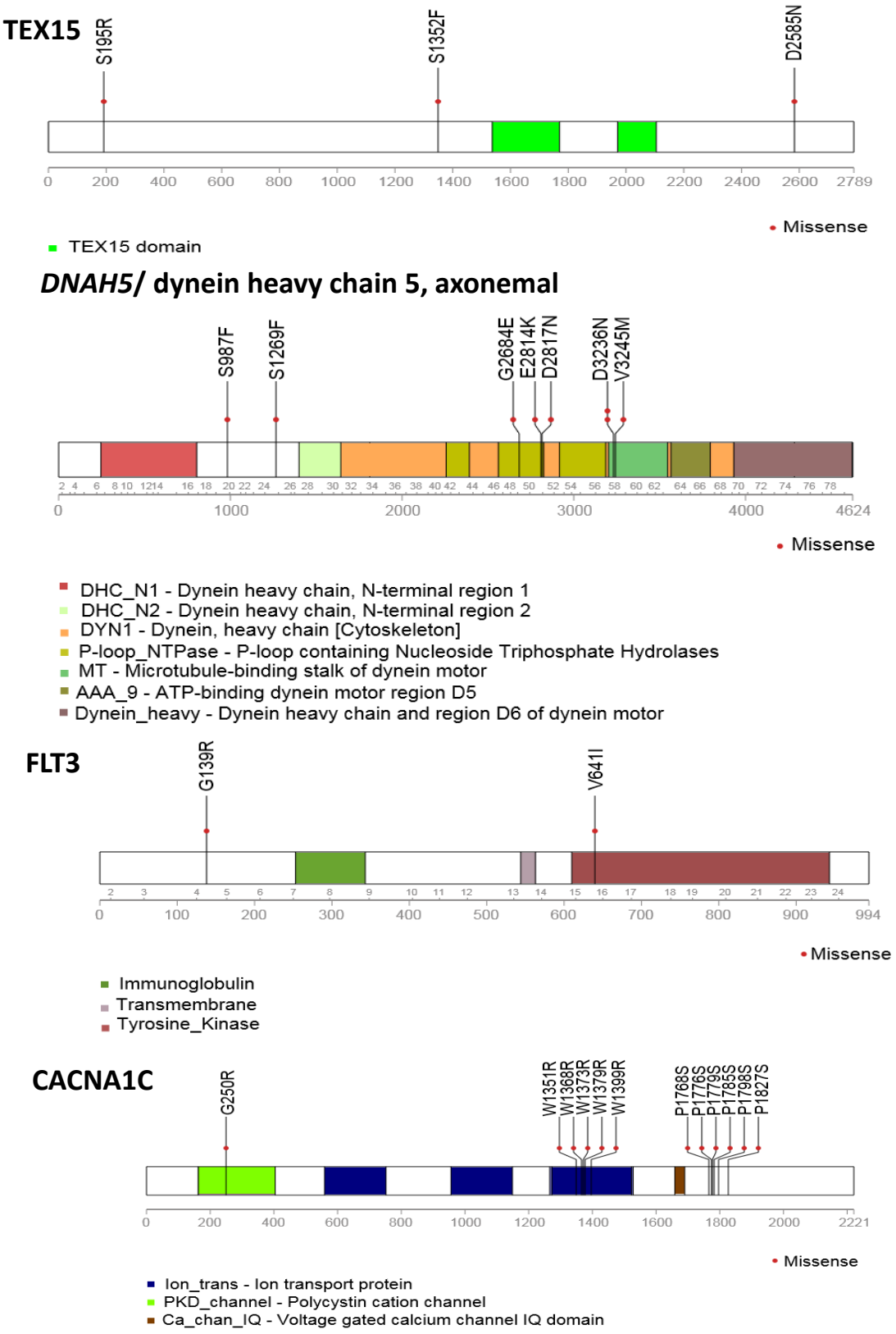


Figure 3.28 (D) Information on the amino acid changes resulting from the mutations in genes selected for target enriched sequencing. The type of somatic mutations (missense, nonsense, frameshift and deletion) in the WES AK data is also shown. Protein names are as per (NCBI) (www.ncbi.nlm.nih.gov/protein/). The diagram was generated using Pediatric Cancer Genome Project (www.explore.pediatriccancergenomeproject.org/proteinPainter).

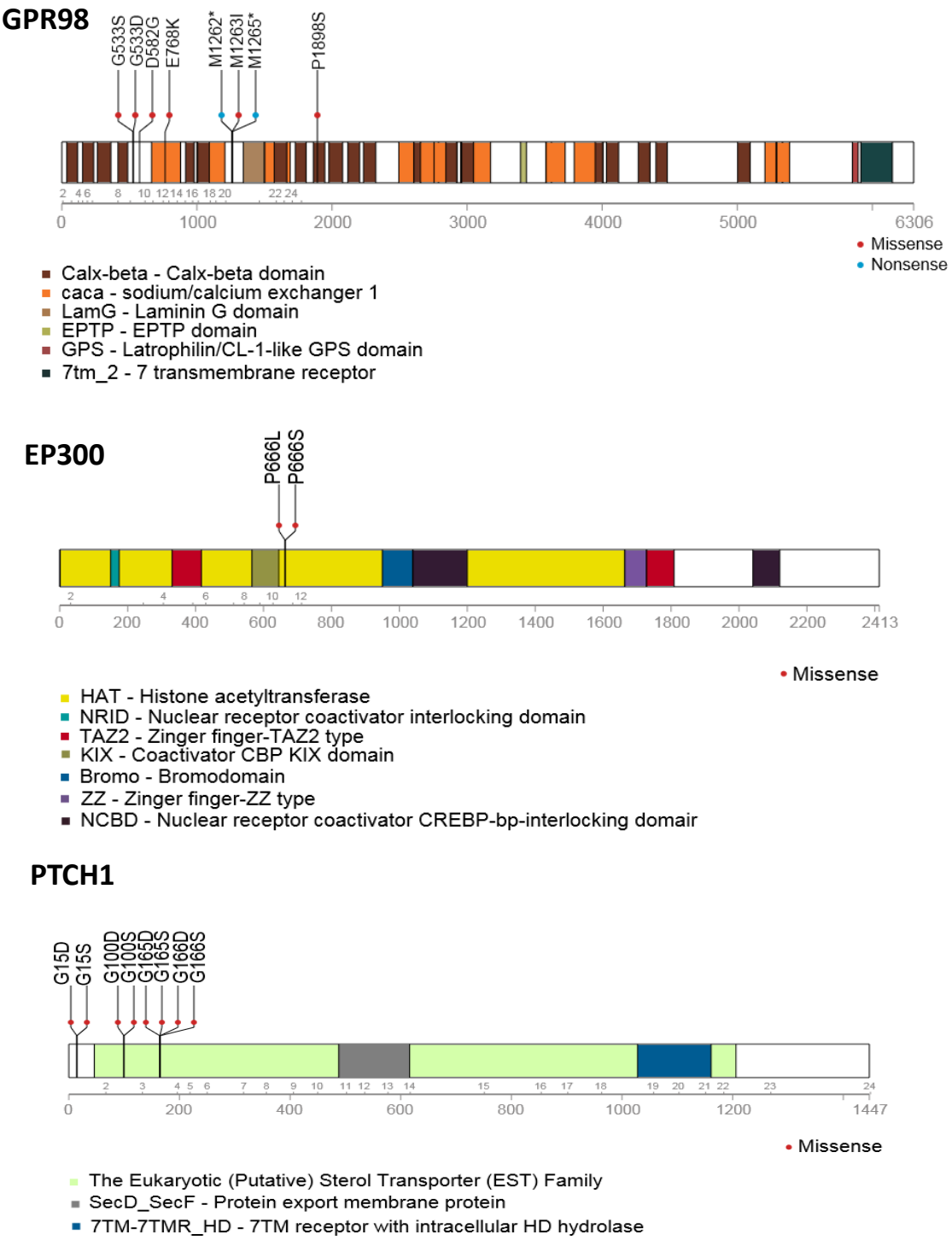


Figure 3.28 (E) Information on the amino acid changes resulting from the mutations in genes selected for target enriched sequencing. The type of somatic mutations (missense, nonsense, frameshift and deletion) in the WES AK data is also shown. Protein names are as per (NCBI) (www.ncbi.nlm.nih.gov/protein/). The diagram was generated using Pediatric Cancer Genome Project (www.explore.pediatriccancergenomeproject.org/proteinPainter).

3.5 Discussion

Various studies have been conducted previously to investigate the molecular defects within AKs and the rate of progression of AK lesions to cSCC. These include retrospective studies on patients who had developed cSCC and observational studies of AK lesions over time examined for malignant transformation (Fuchs and Marmur, 2007, Marks et al., 1988, Quaedvlieg et al., 2006). Investigations of AK have also included parameters such as protein expression (Brasanac et al., 2005, Neto et al., 2013), immune infiltrate (Jang, 2008, Lai et al., 2016) and genetic alterations (Rehman et al., 1994). Genetic studies looking at mutations in AK to date are limited to a few genes including *TP53* (Nelson et al., 1994, Nindl et al., 2007, Park et al., 1996, Soufir et al., 1999, Taguchi et al., 1998, Ziegler et al., 1994) and *CDKN2A* (Nindl et al., 2007, Pacifico et al., 2008, Soufir et al., 1999), and *RAS* gene family (Nindl et al., 2007, Taguchi et al., 1998, Zaravinos et al., 2010). The aim of this chapter was to undertake WES on a small group of AKs to expand the knowledge on mutations in this type of lesion and as a precursor to targeted sequencing of selected genes in a larger group of precancerous lesions. WES studies on NMSC lesions have been conducted in recent years and showed a high mutation burden within those lesions (Bonilla et al., 2016, Durinck et al., 2011, Jayaraman et al., 2014, Li et al., 2015, Pickering et al., 2014, South et al., 2014), thus it was also interesting to know whether AKs contained a high mutation burden because these lesions can develop into NMSC.

As for any genetic studies, targeting the right tissue is important for reliable results. Thus as preliminary step, it was important to characterise the dysplastic area within AK samples. H&E staining for recognition and grading of dysplasia and immunostaining for p53, but not for beta catenin, was useful to differentiate the abnormal keratinocytes from normal epidermis. In addition, staining for the dermal immune infiltrate was helpful to distinguish basal epidermis from underlying immune cells, especially in those samples where the tissue sections were angled (i.e. slightly cross-cut) rather than completely vertical through the epidermis and dermis. Utilising the knowledge gain through this work, it ensured a high purity of targeted tissue when using LCM to dissect the dysplastic and normal skin areas. Likewise, the assessment of DNA quantity, purity, and of the ability to amplify the DNA using PCR, enabled recovery of DNA with sufficient quality and quantity from the target tissue to be used in WES.

Despite the classification of AK as a precancerous lesion, the WES results demonstrated a massive burden of mutation within AKs far beyond that of many other cancers, many of which have a substantial tendency to metastasise. For example, the mutation burden of AKs was ~ 4 fold greater than lung SCC (8.2 mutations/megabase) (Kandoth et al., 2013), 11 fold greater than head and neck SCC (3.2 mutations/megabase) (Kandoth et al., 2013) and ~ 2.3 fold greater than melanoma (16.8 and 14.4 mutations/megabase) (Akbani et al., 2015, Hodis et al., 2012) respectively. By comparing the median number of nonsynonymous mutations across different cancers with high fatality and metastatic tendency, AK lies at the upper part of the distribution in parallel with cSCC as shown in appendix figure 8.1. When the AK mutation burden is compared with that in NMSC lesions, the median mutation rate of 34.5 per megabase within AKs is found to be similar to an earlier study of 8 cSCCs, which recognised a median mutation rate of 33 per megabase (Durinck et al., 2011). However, a more recent study on 20 sporadic cSCC by South et al. (2014) detected a higher median mutation rate of 50 per megabase. The mutation burden in BCC, at 75.8 % mutations per megabase, is around 2 fold higher than that in AK (Jayaraman et al., 2014). The high mutation burden within AK,s as well as in BCCs (Jayaraman et al., 2014) and cSCCs (Li et al., 2015, Pickering et al., 2014, South et al., 2014), suggests that either epidermal keratinocytes have a higher resistance to transformation or that skin has a powerful tumour suppressive activity that stops the clonal expansion of cells harbouring driver mutations.

Despite the high mutation rate detected using WES in the dysplastic keratinocytes from AKs, it was possible to generate a list of mutated genes which were likely to play a role in the growth / development of these lesions and of cSCCs as listed in appendix table 8.13. This is not surprising because genes mutated in cSCC and considered relevant to their development were amongst the criteria used to highlight genes that might be relevant in development of AKs; for example*, TP53, CDKN2A, NOTCH1, NOTCH2* and *HRAS* which were included in this list had been previously reported as mutated in cSCC (South et al., 2014), aggressive cSCC (Pickering et al., 2014) and in metastatic cSCC (Li et al., 2015). However, the fact that these genes were mutated in both lesions would support their role in shared biological events, possibly at an earlier stage in the development of both lesions. In addition, *FAT1* was mutated in AK (current project), cSCC (South et al., 2014), and aggressive cSCC (Pickering et al., 2014), and *CACNA1C* was mutated in both AK (current

project) and cSCC (South et al., 2014). Interestingly, recent genetic analysis of BCC also reported mutations in *NOTCH1* and *NOTCH2*, and *TP53*, as well as mutations in *PTCH1* (which was mutated in up to 75% of the BCCs, with 70% of the mutations deleterious) (Bonilla et al., 2016, Jayaraman et al., 2014).

In two studies of cSCC, *PTCH1* was not among the most frequently mutated genes (Pickering et al., 2014, South et al., 2014) suggesting that *PTCH1* is not a driver gene in cSCC. However, in this WES study of AK, 2 out of the 5 (40%) AK samples showed non-silent mutations within *PTCH1*. In addition, previous work based on clinical observation of AK lesions over time (6 months – 6 years) suggested that 36% of BCC lesions developed on lesions that formerly were diagnosed clinically as AKs (Criscione et al., 2009). Taking these two observations in account, *PTCH1* was included within the list of genes selected for target enriched resequencing.

The high frequency of C to T changes in the WES data supports the view that UVR is the main environmental cause of AK (Brash et al., 1991). However, these UV induced alterations are likely to be present in passenger as well as in driver genes. Distinguishing driver from passenger mutations can be difficult, and this was the case with the AK WES results due to the high mutation load. Hodis et al. (2012) reported a high frequency of non-silent: synonymous somatic mutations in WES analysis of melanoma (2:1). The WES data of the current study showed that the ratio of non-silent: synonymous genetic alterations was 4:1 within AKs and the18 genes selected for targeted sequencing in a larger group of precancerous lesions (chapters 4 and 5) showed high non-silent: synonymous mutations ratio (table 3.8). The overall non-silent: synonymous genetic alteration ratio in normal looking skin adjacent to AKs was also around 4:1. Moreover, all of the following well-known cancer genes (*NOTCH1 and 2, NRAS, HRAS, TP53, KRAS, CDKN2A* gene) and most of the 18 selected genes in table 3.8 that might be expected to play a role in cSCC development were not observed to be mutated in normal skin; this may be due to the relatively low read depth (20X) that was used for the WES study.

Table 3.8 The non-silent: synonymous genetic alteration ratio in the 18 selected genes within the dysplastic area of AKs and adjacent normal looking skin.

Gene ID	Non-silent: synonymous mutations ratio in dysplastic area of AK	Non-silent: synonymous genetic alteration ratio in adjacent normal skin
TP53	17:5	0
NOTCH1	2:0	0
NOTCH2	7:2	0
CDKN2A	2:0	0
HRAS	4:0	0
KRAS	1:0	0
NRAS	1:0	0
CACNA1C	12:3	0
FAT1	9:4	1:0
PAPPA2	3:2	1:0
MLL2	5:0	0:1
DNAH5	7:4	4:3
GPR98	8:3	7:2
HMCN1	4:0	1:0
FLT3	2:1	0
EP300	2:0	0
PTCH1	8:0	1:0
TEX15	3:0	0

Previous studies identified *TP53* mutations within AK ranging from 28% (Park et al., 1996) to 53 % of cases (Nelson et al., 1994). This study showed that *TP53* was mutated in all AK samples (figure 3.27), and supports the view that *TP53* alteration is an early event that precedes substantial expansion of abnormal keratinocytes in the development of cSCC (De Gruijl and Rebel, 2008, Jonason et al., 1996, Rebel et al., 2001, Ren et al., 1997, Wikonkal and Brash, 1999).

NOTCH receptor mutations have also been proposed as a main tumour suppressor mechanism in development of cSCCs (South et al., 2014, Wang et al., 2011). The frequent mutation of *NOTCH* genes within the AKs in the current study (40% for *NOTCH1*, 60% for *NOTCH2*) is in keeping with the report on target enriched sequencing of 27 squamoproliferative lesions which showed *NOTCH1* gene mutation in 49% and *NOTCH2* mutation in 23% of cases (South et al., 2014).

Interestingly, 2 out of the 5 AK samples within this study showed *HRAS* mutation, and one AK with a *HRAS* mutation also contained *NRAS* and *KRAS* mutations. Although mutation of *RAS* occurs less frequently than *TP53* and *NOTCH* in cSCC (Oberholzer et al., 2012, South et al., 2014, Pickering et al., 2014), activated *RAS* mutations have been reported in sun exposed normal skin (Martincorena et al., 2015) consistent with *RAS* mutation being an early event in development of some AKs and cSCCs.

Another well-known mutated gene in cSCC is *CDKN2A*, which was mutated in one of the AKs in the current study and which has been reported in 48% and 23% in metastatic and sporadic cSCCs respectively (Li et al., 2015, South et al., 2014). There have been a few studies to date analysing *CDKN2A* gene mutation within AK, with this gene reported as mutated in 2.7% (Nindl et al., 2007), 50% (Pacifico et al., 2008) and 100% (Soufir et al., 1999) of AKs , and in 5% of 27 squamoproliferative lesions (South et al., 2014). The current WES results of the AKs showed other mutated genes that have also recently been found to be mutated in NMSC as shown in appendix table 8.13. These include *GPR98, CACNA1C* and *MLL2* genes, but the role of these newly discovered genes in cSCC and/or AK pathology have yet to be investigated.

The results of the WES data on 9 AKs that published recently by Chitsazzadeh et al. (2016), in which they reported the mutation burden for 7 AKs, showed an average number of mutations equal to 1,186 variants (range 290–1,873) and average mutation burden of 18.5 per megabase (lower than the mutation burden/megabase of AKs from the current study). Consistent with our results, 5 out of the 13 frequently mutated genes identified by Chitsazzadeh et al. (2016) were also identified as mutated within the AK WES data. The frequently mutated genes list in common between the current study and that of Chitsazzadeh et al. (2016) includes *TP53, NOTCH1, NOTCH2, FAT1 and MLL2.* Moreover, they identified *KNSTRN* mutations in only 1 AK sample which was similar to my finding of no *KNSTRN* gene mutations within the 5 AKs investigated by WES. However there were also some differences between the results of the two studies, for example *FLT3* was within the top 18 selected genes in the current project but was not identified as a frequently mutated gene by Chitsazzadeh et al. (2016), while *FAM135B* gene was mutated in 2 out of the 7 AKs by Chitsazzadeh et al. (2016) study but it was not mutated in the current study. Indeed, due to the high mutation burden within AKs, it is also not surprising that the frequently mutated gene list in the current study and the Chitsazzadeh et al, (2016) study differs slightly from the list of 18 genes selected for further study in this thesis. Although the WES data from the current study was informative as to the mutation burden in AK, the small size of the study (N = 5 AKs), meant that it would be sensible to undertake sequencing on the selected genes in larger group of AKs as well as in other precancerous lesions such as Bowen's disease to better characterise the relationship between such lesions and cSCC.

Chapter 4: Target enriched sequencing of Actinic Keratoses (AKs) and Bowen's Disease (BD)

4.1 Introduction

The results of WES analysis of Actinic Keratoses (AKs), described in chapter 3, together with published data describing genes mutated in cSCC (published in Durinck et al. (2011), Li et al. (2015), Pickering et al. (2014), South et al. (2014)) and the COSMIC database, led to the identification of a list of 18 genes that were likely to be frequently mutated in precancerous skin lesions, including AK and Bowen's Disease (BD). To determine whether this is the case, sequencing of the relevant genes in a greater number of precancerous lesions is necessary. Clinical observational studies have indicated that cSCCs can develop from AKs (Criscione et al., 2009, Marks et al., 1988) and BD (Cox et al., 1999, Kao, 1986). Thus, studying these genetic changes within AK and BD lesions may help in establishing the timing of these genetic alterations in the cSCC development process.

Previous studies have detected point mutations in *TP53* and *CDKN2A* within cSCC, AK and BD lesions (Campbell et al., 1993b, Nelson et al., 1994, Nindl et al., 2007, Pacifico et al., 2008, Park et al., 1996, Soufir et al., 1999, Taguchi et al., 1998, Takata et al., 1997, Ziegler et al., 1994). WES and target enriched sequencing studies on cSCC have shown a high mutation burden within cSCC (Durinck et al., 2011, Li et al., 2015, Pickering et al., 2014, South et al., 2014) which leads to challenges in identifying which of these are driver mutations within the cSCC genome. It has been suggested that mutations in certain genes are likely to be driver mutations because the gene is affected in a high frequency of cSCCs and/or there is evidence in the literature that the relevant gene and/or mutation affects cell behaviour in a way that might lead to cancer development. Thus, to date, potential driver genes as indicated in the cSCC genetic studies include the *NOTCH* family genes (particularly *NOTCH1& NOTCH2*) (Durinck et al., 2011, Li et al., 2015, Pickering et al., 2014, South et al., 2014), *TP53* (Brash et al., 1991, Kubo et al., 1994, Tabata et al., 1999, Ziegler et al., 1994)*, CDKN2A* (Brown et al., 2004, Saridaki et al., 2003)*, RAS* family genes (Oberholzer et al., 2012, Su et al., 2012) and may be *PTCH1* (Ping et al., 2001). The aim of the current study was to sequence each of these genes (as well as others listed in the group of 18 genes chosen for target enriched sequencing in chapter 3) in AKs and BDs to provide additional support for these genes being involved in development of cSCC.

Recently Lee et al., 2014 reported mutation of the *KNSTRN* gene within AK and cSCC lesions. *KNSTRN* encodes a kinetochore-associated protein that controls the beginning of

anaphase and chromosome segregation during mitosis (Dunsch et al., 2011). It is expressed in a variety of human tissues, including skin ((The Human Protein Atlas) (www.proteinatlas.org/), (Berglund et al., 2008, Uhlén et al., 2005)). Although *KNSTRN* had not been identified as being within the top mutated genes in WES studies of cSCC (Durinck et al., 2011, Li et al., 2015, Pickering et al., 2014, South et al., 2014) and AK (chapter 3 in this thesis), point mutations within the *KNSTRN* gene were noted by Lee et al., 2014 in 19% of cSCCs and 17% of AKs. The frequency of the detected alterations (p.Arg11Lys, p.Ser24Phe, p.Pro26Ser, p.Pro28Ser and p.Ala40Glu) varied, however all documented mutations were distributed in the first exon of the gene with the p.Ser24Phe amino acid change the most frequent (Lee et al., 2014). For this reason, it was decided to conduct additional Sanger sequencing of exon 1 of the *KNSTRN* gene on the AK and BD samples.

4.2 Hypothesis and aims

The hypothesis was that genes in which mutations are thought to be potential driver gene mutations in cSCC may also be mutated in AKs and BD. The aim of this chapter was to undertake target enriched sequencing of a selection of 18 genes (chosen because they were mutated in the AK WES data (chapter 3), cSCCs in the published literature and in COSMIC database) in the potential precancerous skin lesions, AK and BD.

4.3 Materials and Methods

4.3.1 Tissue sample

Archived FFPE blocks containing AK and, separately, FFPE blocks containing BD were collected from the Histopathology department at Southampton General Hospital. Blood DNA samples (control) were obtained from patients as part of an investigation on skin cancer being conducted by Dr Chester Lai in Dermatopharmacology, University of Southampton. The study was conducted under local research ethics committee approval and a signed consent form was obtained from all subjects for the use of the tissue.

4.3.2 Haematoxylin & Eosin (H&E) staining

Sections (5 µm thick) of AK, BD, AK adjacent to cSCC and BD adjacent to cSCC were cut from the FFPE blocks and were stained for H&E as described in materials and methods section 2.2.5 to allow identification of the dysplastic area (and cancerous area in the cSCC

sections) in the lesions. For the samples where two lesions were adjacent to each other (cSCC / AK and cSCC /Bowens disease) (N = 15), the dysplastic part of the AK and of the BD and the region of the section containing the cSCC was confirmed by Dr. Jeffrey Theaker, Consultant Histopathologist, Histopathology department, University Hospital Southampton NHS Foundation Trust. Sections with a small area of dysplasia or with a heavy immune infiltrate were excluded because of the difficulty of isolating sufficiently pure samples of dysplastic and/or cancerous cells from these samples.

4.3.3 Laser capture microdissection and cresyl violet acetate staining

To ensure the enrichment of dysplastic lesional and/or cancerous keratinocytes within the samples, 10 µm thick FFPE tissue sections were laser capture microdissected after being stained with cresyl violet acetate as described in section 2.2.8. The matched adjacent non-dysplastic perilesional normal looking skin tissue was also laser microdissected and used as control. The dissected tissue was collected in a microcentrifuge tubes containing 180 µL of ALT lysis buffer for subsequent DNA extraction.

4.3.4 DNA extraction

Genomic DNA was extracted from the lesional (dysplastic and/or cancerous) and matched adjacent normal looking skin tissue which had been microdissected as above (4.3.3). Genomic DNA was purified (as outlined in section 2.2.9) using the QIAamp DNA FFPE Tissue Kit (Qiagen) according to the manufacturer's instructions.

4.3.5 Assessment of DNA concentration within the samples

A Qubit 2.0 fluorometer (Invitrogen) was used for DNA quantification as described in section 2.2.10. According to Illumina, the target DNA concentration for target enriched sequencing should be ≥ 5 ng/µl with total DNA amount of ≥ 100 ng.

4.3.6 Assessment of DNA quality

As described in section 2.2.11, extracted genomic DNA was assessed for purity and fragmentation. Using Nanodrop D-1000 (Nanodrop Technologies, US) absorbance spectrophotometer, the optical density in the form of A260/A280 and A260/A230 ratios was measured for each DNA sample. DNA fragmentation was assessed (as detailed in section 2.2.11) using a panel (multiplex) PCR that amplifies different DNA fragment

lengths (100, 200, 300, 400 and 600 base pairs). The primers and reaction conditions for panel PCR were as detailed in the materials and methods section 2.2.11.1. The resultant bands were visualised on an agarose gel that was prepared as detailed in section 2.2.11.2.

4.3.7 Target enriched sequencing

4.3.7.1 Kit design for customised target enriched sequencing

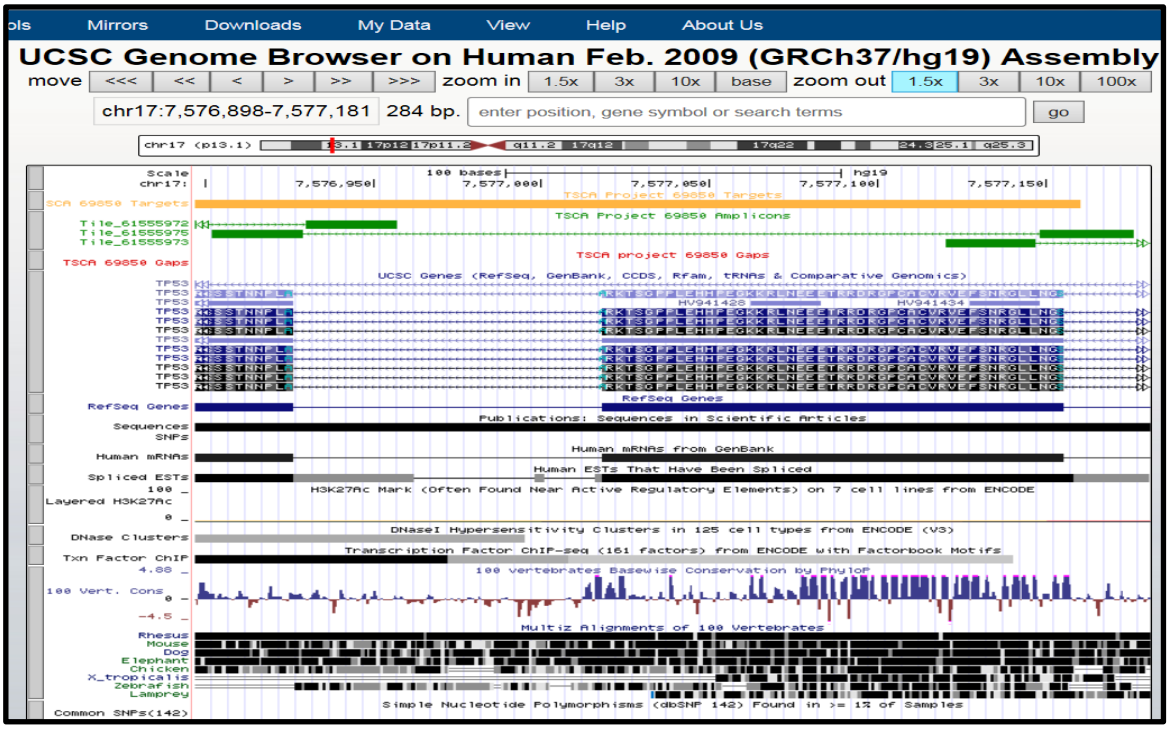
4.3.7.1.1 Selection of the kit for target enriched sequencing

Different companies provide a range of kits that can be customised to target selected genes on the genomic DNA. Several investigations were conducted in the process of selecting the most suitable kit and design for the targeted genes. The factors that were taken in consideration included the cost and the concentration (input amount) of the DNA required by the different kits. As a result, the TruSeq custom amplicon kit v1.5 provided by Illumina was selected.

4.3.7.1.2 TruSeq custom amplicon (TSCA) v1.5 kit design

Primers was designed using the Illumina DesignStudio programme (Illumina) (www.designstudio.illumina.com/) and validated by mapping the design to the GRCh37/hg19 (Genome Reference Consortium 37) of the human genome to ensure appropriate coverage of the targeted DNA region. A total of 1,344 amplicons (each of 250 base pairs in length) representing the exons of the 18 selected genes were targeted during primer design and an example of the process of kit design is shown in figure 4.1. The final total coverage of the design was 98%. Information on the 18 targeted genes, and the exonic area of the genes that could not be covered by the kit primers, is detailed in tables 4.1 & 4.2.

A



B

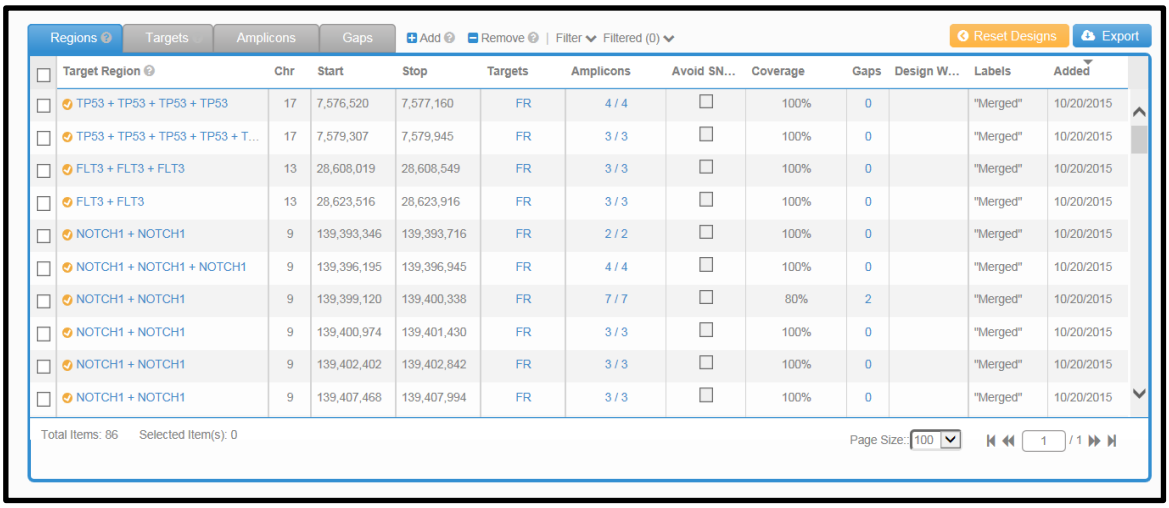


Figure 4.1 Screenshot of the process of designing the TruSeq custom amplicon (TSCA) v1.5 kit. **A**: Screenshot image from the UCSC genome browser (UCSC) (www.genome.ucsc.edu/) shows the process of manipulating the amplicon design for *TP53*, in order to reduce the amplicon size from 284 base pairs to a slightly smaller size of 250 base pairs, which is the selected amplicon size. **B**: Screenshot image for the final design showing some of the targeted areas and the percentage coverage of the target area.

Table 4.1 Percentage of coverage of the exons for 18 genes selected for target enriched sequencing. Details of the chromosome where gene is located, gene length in base pairs, number of exons, length of the transcript and the percentage of the coverage for each gene are indicated in the table. Information about the genes was obtained from (NCBI) (www.ncbi.nlm.nih.gov/gene) and (Ensembl) (www.ensembl.org/).

Gene ID	Chromosome	The start of the gene on the chromosome	The end of the gene on the chromosome	The gene length in base pairs	The number of exons within the gene	The length of the transcript in base pairs	Percentage of gene exon sequence coverage
CACNA1C	12	2,162,416	2,807,115	644,800	47	13,433	100%
CDKN2A	9	21,967,751	21,994,490	26,840	3	1,283	100%
DNAH5	5	13,690,437	13,944,589	254,253	79	15,633	96.9%
EP300	22	41,488,614	41,576,081	87,568	31	9,585	97%
FAT1	4	187,508,937	187,644,987	136,151	27	14,786	100%
FLT3	13	28,577,411	28,674,729	97,419	24	3,842	97%
GPR98	5	89,854,617	90,460,033	605,517	90	19,338	95.1%
HMCN1	1	185,703,683	186,160,085	456,503	107	18,208	100%
HRAS	11	532,242	535,550	3,409	6	894	100%
KRAS	12	25,358,180	25,403,854	45,775	5	5,765	100%
MLL2	12	49,412,758	49,449,107	36,450	54	19,419	95.5%
NOTCH1	9	139,388,896	139,440,238	51,443	34	9,371	97.5%
NOTCH2	1	120,454,176	120,612,317	158,242	34	11,389	100%
NRAS	1	115,247,085	115,259,515	12,531	7	4,449	100%
PAPPA2	1	176,432,307	176,811,970	379,764	23	9,691	96%
PTCH1	9	98,205,264	98,279,247	74,084	24	8,057	100%
TEX15	8	30,689,060	30,706,533	17,574	4	10,187	100%
TP53	17	7,668,402	7,687,538	19,136	11	2,506	100%
Overall coverage for 18 genes							98%

Table 4.2 Gene sequences that are not covered by the TruSeq custom amplicon (TSCA) v1.5 kit. The positions and the lengths of the sequences within the selected genes for which the TruSeq design programme could not provide primers to amplify these regions by PCR for target enriched sequencing (labelled as "gaps" in the table).

Gene ID	Chromosome	Gap start position on the chromosome	Gap end position on the chromosome	Coding protein	Reported to be mutated previously in Ensembl database	Variants as reported in dbSNP Build 137	Reported to be mutated in cSCC/AK	Gap length in base pairs
DNAH5	5	13,691,519	13,691,685	Yes	No		No	166
EP300	22	41,572,246	41,572,331	No				85
FLT3	13	28,597,619	28,597,695	Yes	No		No	76
GPR98	5	90,149,899	90,149,938	Yes	Yes (2 missense variants)	rs758256273 rs727504913	No	39
MLL2	12	49,426,927	49427,030	Yes	Yes (1 missense variants)	rs771142956	No	103
NOTCH1	9	139,390,408	139,390,451	Yes	No		No	43
PAPPA2	1	176,680,966	176,680,981	Yes	No		No	15

4.3.7.2 Target enriched sequencing

Genomic DNA from the lesional and matched adjacent normal looking skin samples from AKs (N = 25), BDs (N = 29), cSCCs adjacent to either AK or BD (N = 15) and adjacent normal looking skin were selected for target sequencing. In brief, between 119 and 2,467 ng of DNA per sample was sent on dry ice to the Wellcome Trust Centre for Human Genetics, University of Oxford for sequencing. The customised TSCA v1.5 kit from Illumina was used for target capture and library preparation and the Illumina MiSeq 2000 platform was used for DNA sequencing according to the manufacturers' instructions as described in section 2.2.12.2.

4.3.7.3 Read mapping and variant calling

The initial bioinformatics analysis including read mapping and variant calling was conducted by Dr Reuben Pengelly, Genetic Epidemiology and Genomic Informatics Group, Human Development and Health, Faculty of Medicine, University of Southampton. Initial analysis involved sequencing reads being mapped to the human GRCh37/hg19 (Genome Reference Consortium 37) using BWA (Burrows-Wheeler Aligner) (Li & Durbin, 2009). Files for called variants were generated for all samples using SAM Tools and lesional/ adjacent normal skin sequencing pairs were used as input for VarScan 2.3.3 (Koboldt et al., 2012) assuming a lesion purity of 90% and requiring a minimum variant frequency of 10%. Variants that were flagged as somatic were retained and passed to ANNOVAR to be annotated. Final output files were then converted to a tabulated excel format and provided for further analysis.

Following initial analysis, filtering steps were implemented to exclude variants that were flagged as homopolymer (highly repetitive regions), and to ensure a minimum mean fold coverage for the targeted genome of ≥ 100x, a p value for somatic changes ≤ 0.05 and selected variants ≥ 4 of the lesional reads and < 4 in normal reads.

4.3.8 *KNSTRN* gene PCR and Sanger sequencing

PCR for *KNSTRN* was conducted as detailed in section 2.2.11.1. The resultant bands were visualised on agarose gels prepared as detailed in section 2.2.11.2. Sanger sequencing was conducted by Source BioScience (Nottingham, UK) and the results (AB1 files) were

analysed using SeqBuilder (v10.0.1.3) and Seq Scanner (v2.0) software's (Applied Biosystems).

4.3.9 Statistics

GraphPad Prism (v6.0) was used to generate figures and to undertake statistical analysis (support was received from Scott Harris and Ho-Ming (Brian) Yuen, Medical Statistics, Primary Care and Population Sciences, Faculty of Medicine, University of Southampton).

4.4 Results

As described in materials and methods section 2.2.2, archived FFPE human samples were identified and collected from Histopathology, University Hospital Southampton NHS Foundation Trust. Details of samples, including gender and age of patient and site of skin lesion are provided in table 4.3 and appendix table 8.15.

Table 4.3 Details of gender and age of patients and site of skin lesions used for target enriched sequencing.

Nature of the lesion	**Solitary AK**		**Solitary BD**		**AKs adjacent to cSCCs**		**BDs adjacent to cSCCs**	
Gender	Male	Female	Male	Female	Male	Female	Male	Female
	19	6	17	12	6	1	8	0
Site of the lesion								
Scalp	7	1	7	2	3	0	4	0
Cheek	2	1	5	1	1		2	
Ear	3		3					
Nose						1	2	0
Hand	3	1	2	2	1			
Forearm/arm	0	1			1			
Neck	2	1		2				
Chest/back	1	1						
Lower leg	1			5				
Total Number of samples	25		29		7		8	
Age at biopsy	Range (54 - 92years), Median 74 years		Range (57 – 94years), Median 84 years		Range (65 - 88years), Median 81 years		Range (57 - 94years), Median 80 years	

H&E staining was conducted on all samples in order to allow identification of the dysplastic and/or cancerous area of the lesion. Examples of H&E staining are shown in figure 4.2.

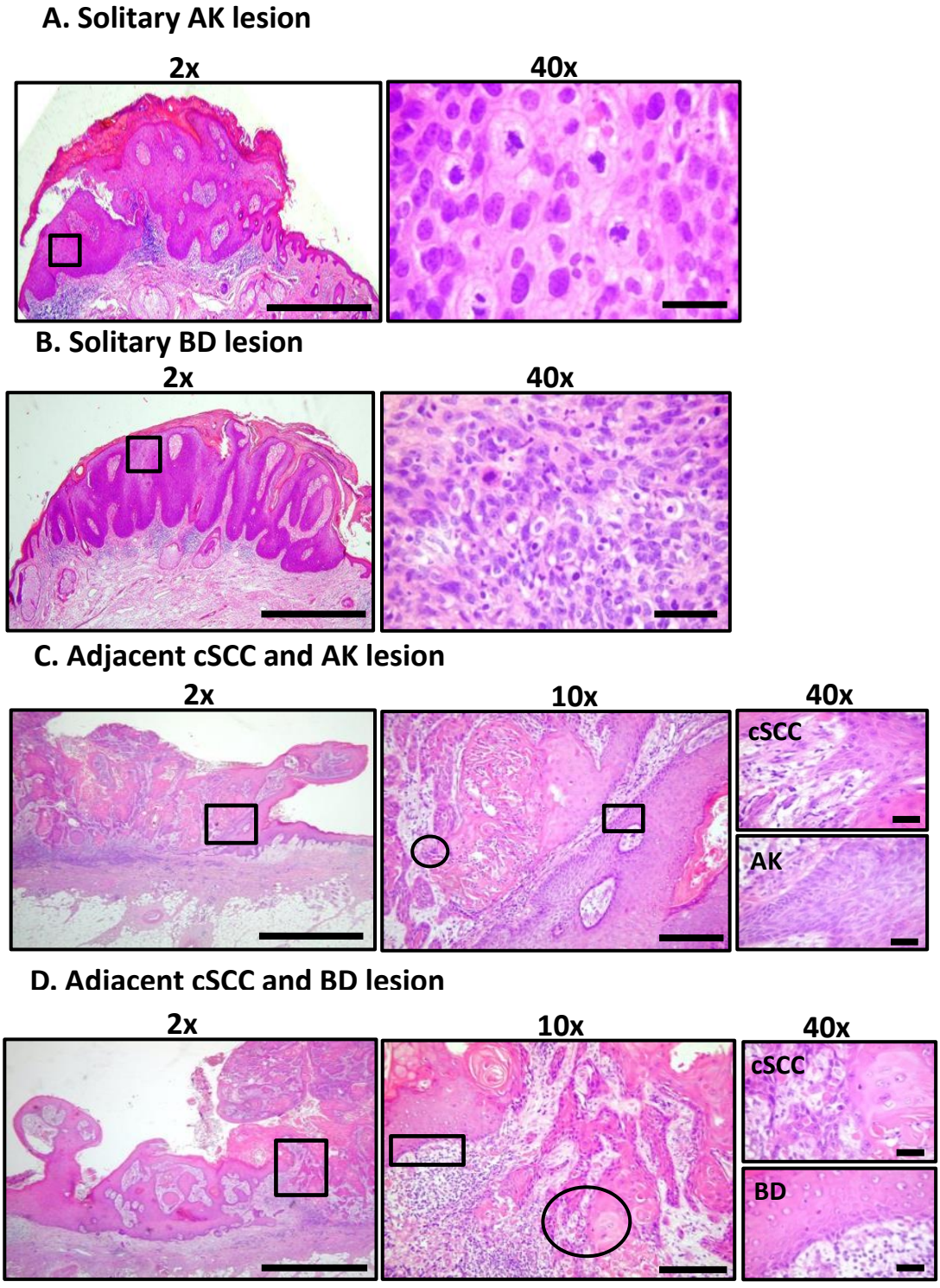


Figure 4.2 H&E staining of the different skin lesion types selected for target enriched sequencing. **A&B:** images (at 2x and 40x) represent solitary AK and solitary BD lesions respectively. **C:** Adjacent cSCC and AK lesion. **D:** Adjacent cSCC and BD lesion. For each lesion, 3 images at different fold magnifications were taken (2x, 10x and 40x). Black boxes in images 2x of C & D represents the area of 10x images. In the 10x images, the boxes represent the AK (in **C**) and BD (in **D**) lesions within the 40x images respectively, while the black circles represent the area of cSCC within the 40x images. AK = Actinic keratosis. BD = Bowens disease. cSCC = Cutaneous squamous cell carcinoma. Scale bar; 500, 50 & 10 µm for 2x, 10x, and 40x magnifications respectively.

Following examination of the H&E stained sections, a total of 69 different skin samples were selected for target enriched sequencing (as listed in table 4.3). These comprised 32 AKs (including 7 AK/cSCCs where a cSCC was adjoining an AK in the same histological section and considered to have arisen from the adjacent AK) and 37 BDs (including 8 BD/cSCCs where the cSCC was adjoining BD and seemed to have developed from the BD) as well as corresponding non-lesional skin. The dysplastic and/or cancerous area of the lesions were microdissected using laser capture microdissection and, following lysis of these tissue samples, DNA was extracted from a total of 153 tissue samples that represented the lesional and/or cancerous areas and matched non-lesional adjacent normal looking skin. The total DNA yield ranged from 119 – 2,467 ng, and the DNA quantity for the individual lesions are summarised in figure 4.3. In general, as expected BD samples gave more DNA than the smaller AK lesions (average total amount of DNA was 385.2 ng for AKs and 425.6 for BDs) (appendix table 8.16). A Nanodrop spectrophotometer was used to assess DNA purity and the results of the 230/280 and 260/280 ratios are summarised in figure 4.4. In addition, DNA fragmentation assessment was also conducted. The results showed that 18 of the 32 AKs, 19 of the 37 BDs and 7 of the 15 cSCCs showed DNA fragment length equal to 300 bps. Moreover, 10 of the 32 AKs, 11 of the 37 BDs and 7 of the 15 cSCCs showed DNA fragment length equal to 400 bps and 4 of the 32 AKs, 7 of the 37 BDs 1 of the 15 cSCCs showed DNA fragment length equal to 600 bps.

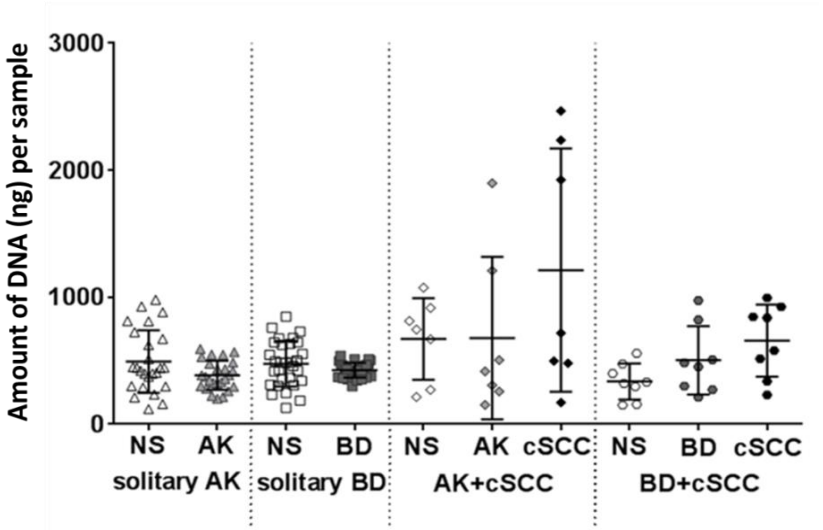


Figure 4.3 DNA yield for the 69 lesions selected for target enriched sequencing. DNA amount was measured using a Qubit 2.0 fluorometer and expressed in ng. AK = Actinic keratosis. BD = Bowens disease. cSCC = Cutaneous squamous cell carcinoma. NS = Normal skin.

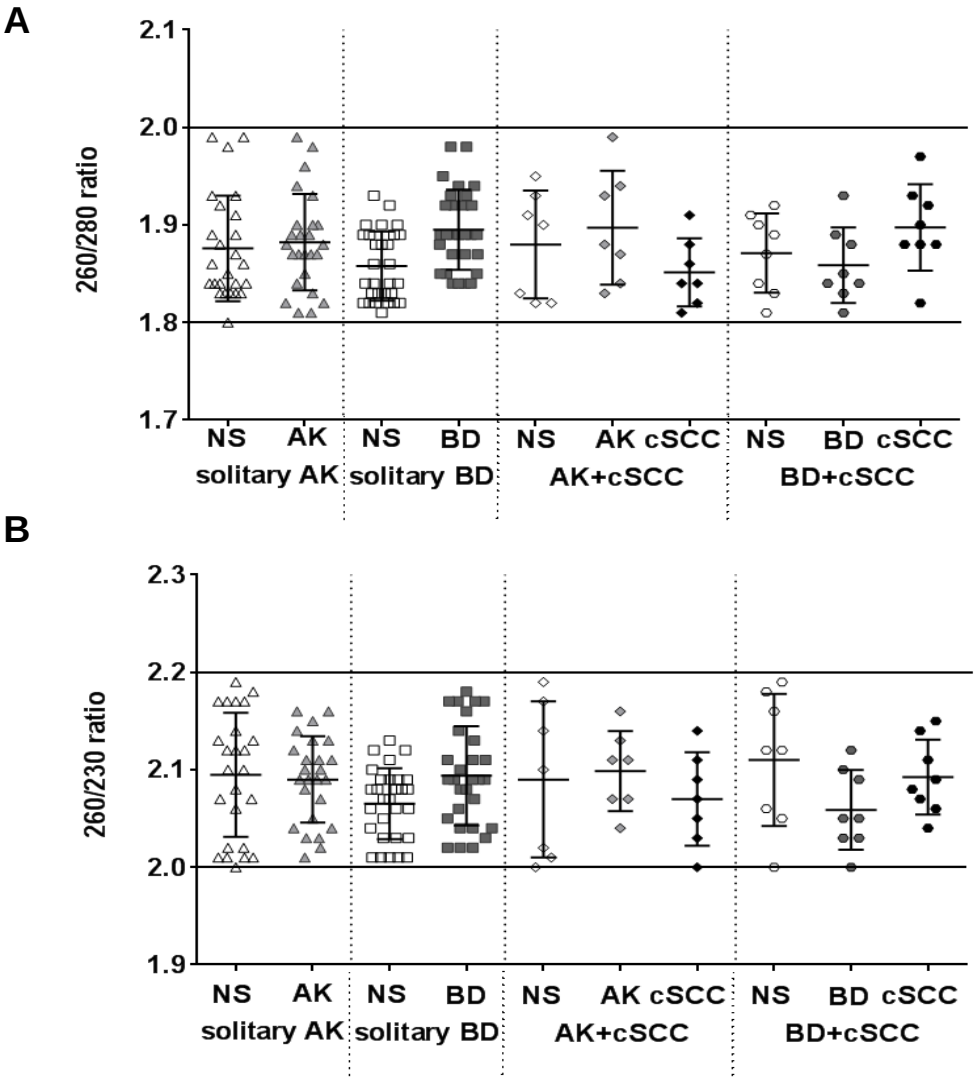


Figure 4.4 Nanodrop spectrophotometer quality assessment of DNA extracted from the 69 FFPE lesions selected for target enriched sequencing. **A:** A260/A280 ratio (values between 1.8 and 2.0 indicate high purity DNA with minimal protein contamination). **B:** A260/A230 ratio (values between 2.0 and 2.2 indicate high purity DNA with minimal solvent contamination). AK = Actinic keratosis. BD = Bowens disease. cSCC = Cutaneous squamous cell carcinoma. NS = Normal skin.

4.4.1 Target enriched sequencing data

4.4.1.1 Primary data analysis

Target enriched sequencing was conducted on DNA isolated from the 69 selected FFPE skin samples and matched adjacent normal looking skin (25 AKs, 29 BDs & 15 cSCCs adjacent to either AK [7 samples] or BD [8 samples]) using a TruSeq custom ampliconv1.5 kit designed to target the 18 genes selected according to the criteria discussed in section 3.4.5.6. The initial results showed that the mean depth of coverage varied across the different lesions. Bowen's disease samples had the lowest mean depth of coverage at 166 fold and cSCC samples had the highest mean depth of coverage at 205 fold. Only somatic

non-silent (missense, nonsense, frameshift insertion/deletion and splicing) alterations were included within the analysis. The median number of mutations per sample ranged from 10 – 25.5. However, there was no consistent pattern of median (or mean) number of mutations across the dysplastic lesions and cancerous lesions. The results are illustrated in table 4.4.

The total size of the 18 genes (gene's exons) that were sequenced using target enriched sequencing was 196 kilobases, therefore the mutation burden for target enriched sequencing is expressed in kilobase as well as in (the more conventional) megabases. It should be noted that the use of megabase generates a high mutation burden for many of the lesions (for example the median mutation burden for solitary AKs was 130 mutation /megabase of DNA), but this is can be explained by the fact that the 18 genes were selected because they were frequently mutated in AK and cSCC.

Table 4.4 Target enriched sequencing primary data results. AK = Actinic keratosis. BD = Bowen's disease. cSCC = Cutaneous squamous cell carcinoma. Kilobase (Kb). Megabase (Mb).

	Solitary AK	**Solitary BD**	**AK adjacent to cSCC**	**BD adjacent to cSCC**	**cSCC adjacent to AK**	**cSCC adjacent to BD**
Number of samples	25	29	7	8	7	8
Mean fold coverage	179.5 (range 79.7 – 305.4)	166 (range 0.55 – 369.7)	191.4 (range 40.8 – 302.9)	183.6 (range 37.1 – 274.5)	205 (range 72.5 – 313.2)	188.5 (range 51.1 – 269.4)
Total number of non-silent mutations	1345	759	256	530	260	517
Median number of mutations /sample	25.5	10	14	24	13	16
Mean number of mutations /sample	56.04	27.1	36.5	66.5	37.1	64.6
Range of mutations /sample	6-296	1-131	2-149	9-216	12-124	3-225
Median mutation burden /Kb of DNA	0.13	0.05	0.07	0.12	0.07	0.08
Median mutation burden /Mb of DNA	130	50	70	120	70	80

The mean numbers of mutations per each of the tested genes were compared between the lesions. The overall mean numbers of mutations were high in genes such as *MLLS, GPR98* and *HMCN1* genes and low in other such as *TP53, CDKN2A*, and *RAS* family genes. This seemed inconsistent with the previous published results from research on *TP53* (Nelson et al., 1994, Nindl et al., 2007, Pacifico et al., 2008, Park et al., 1996, Soufir et al., 1999, Taguchi et al., 1998, Takata et al., 1997, Ziegler et al., 1994) and indicated that, for

more informative results, additional filtration steps were required. The results are shown in figure 4.5.

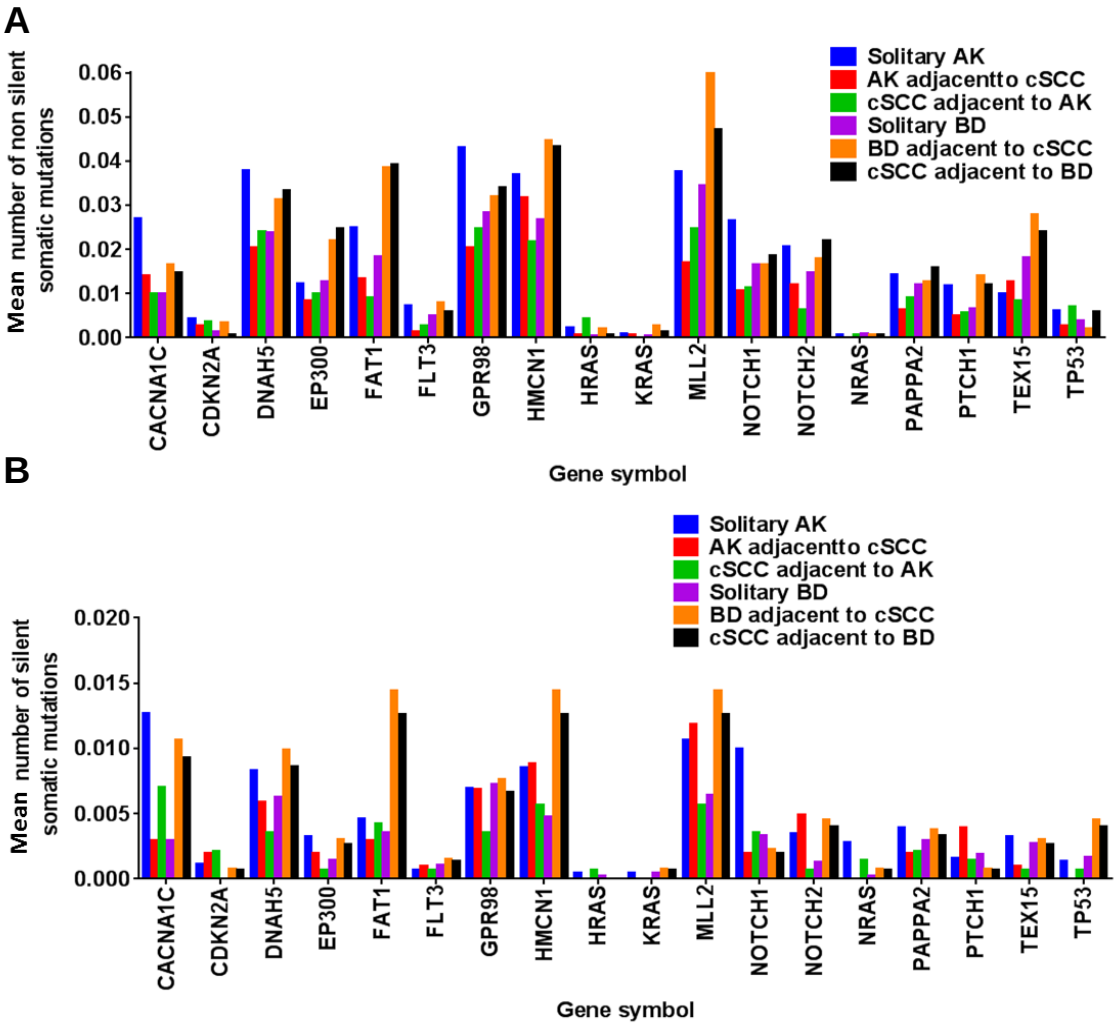


Figure 4.5 Overview of the primary analysis of the target enriched sequencing data for the individual genes. *MLL2, GPR98* & *HMCN1* showed a high number of mutations within most of the lesions. **A:** non-silent mutations. **B:** silent mutations.

4.4.1.2 Variants call after the second set of filtration steps

Similar to other NGS studies in the literature (Durinck et al., 2011, Li et al., 2015, Pickering et al., 2014, South et al., 2014), additional filtration steps were applied to the data, thus removing mutations which did not fulfil the filtration criteria. These filteration criteria were selected according to the target enriched sequencing analysis from studies on cSCC (Li et al., 2015, Pickering et al., 2014, South et al., 2014) and included a minimum mean fold coverage of ≥ 100x, variants denoted as mutations being present in ≥ 4 reads from the lesion and < 4 reads from the normal skin, and a p value for somatic changes of ≤ 0.05 (as described in section 4.3.7.3). While the second filtration step reduced the numbers of

mutations and lesions from each group in the final set of results, it made the resulting data more robust.

4.4.1.2.1 Solitary AK data analysis

Out of the 25 solitary AK samples, 18 AKs passed the second filtration run, leading to a total number of non-silent somatic genetic alterations of 308 (out of 1345 alterations before the second filtration step) with a median of 10 mutations per lesion, range 3 – 56. The ratio of non-silent to silent (synonymous) somatic mutation was 3.7:1. Consistent with the fact that UVR exposure is the main risk factor for development of AKs (Tsatsou et al., 2012), C > T transition base substitutions were the most common type of base pair change (median frequency of 61.2%, range 28.5 – 100%). The overall frequency of CC > TT was 7% and was seen in only 11 of the 18 AK samples. Although, after the second filtration, *MLL2* remained within the list of genes with a high mutation frequency (equal to 78%), several other genes, including *CANA1C* (78%), *NOTCH1* (72%), and *NOTCH2* (61%) genes were also commonly mutated. Most of the mutations (as seen in the heat map in figure 4.6) were missense mutations, with *TP53* showing the highest number of truncating mutations (5 out of 12 mutations).

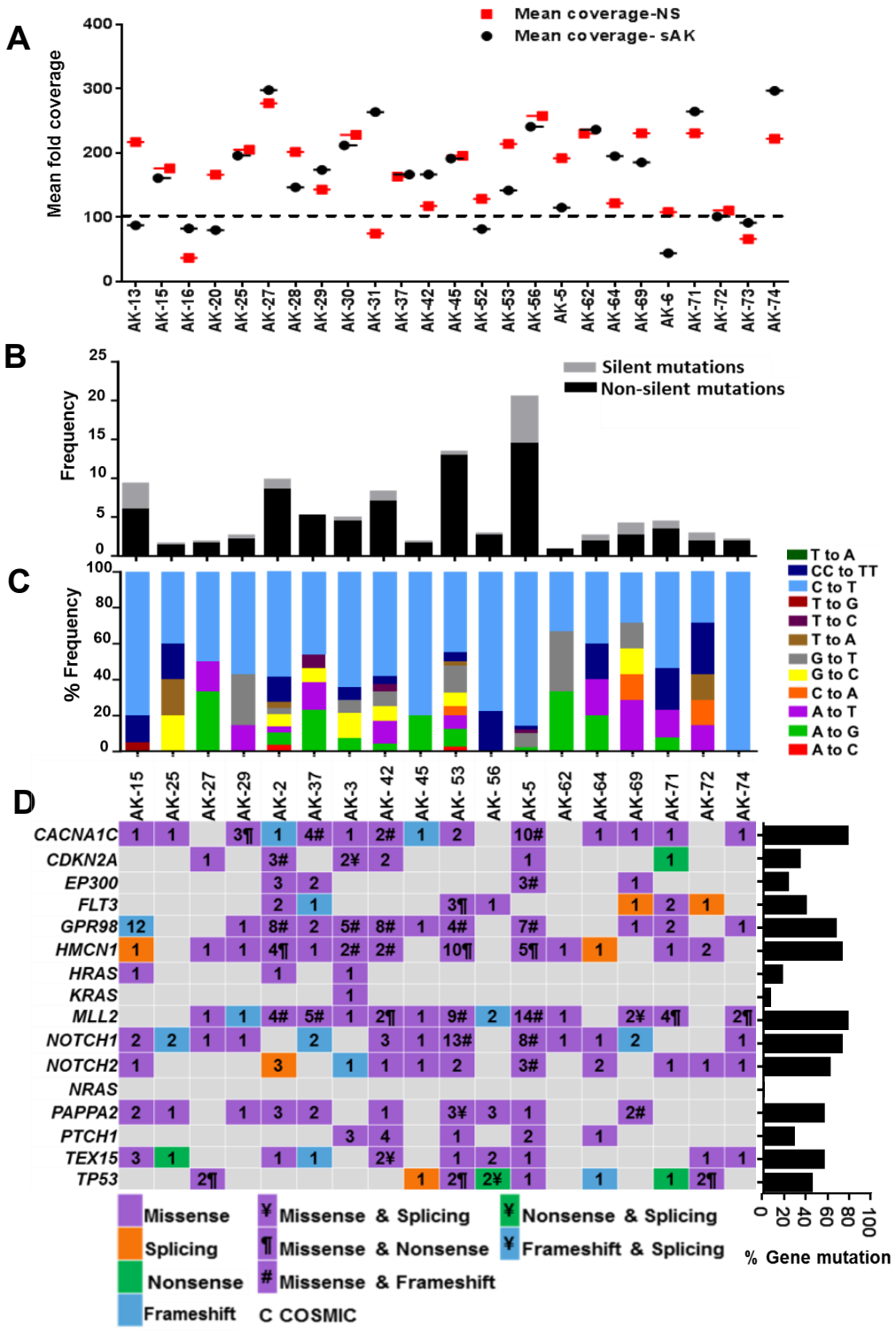


Figure 4.6 Integrated overview on the target enriched sequencing data from solitary AK (sAK); part A shows results from the 25 sAKs which underwent target enriched sequencing, whereas parts B – D show results from the 18 sAKs remaining after the second round of filtration. **A:** The mean fold coverage for the 18 genes across the 25 samples within the AK study. **B:** The frequency of silent and non-silent genetic alterations within AKs. **C:** The frequency of the different base pair changes within the AKs showing that C > T changes were the most common. **D:** Heatmap representation of the different types of mutations within the target genes for each of the AK samples. The numbers within the heatmap indicate the number of mutation/gene/sample with the different colours each indicating a different type of mutation. The corresponding bar graph on the right of the heatmap represents the % of gene mutations within targeted AKs. Note that in B – D, each of the columns represent the same AK sample.

4.4.1.2.2 Solitary Bowen's disease data analysis

After the second filtration step, 17 samples out of the 29 solitary BD samples remained and were included in the analysis. Mutations were detected in all of these BD samples, and the number of somatic mutations per BD ranged from 2 – 24 (median = 7). The ratio of non-silent somatic mutations to silent (synonymous) mutation was 4:1 in these solitary BDs. C > T base substitutions, suggestive of UVR-induced mutations were frequent (median frequency of 58%, range 0 – 100%), with CC > TT mutations reported in 6 of the 17 BD samples at overall frequency equal to 9%. Missense mutations were the most common type of non-silent mutations (as shown in figure 4.7). *NOTCH1* was the most mutated gene, occurring in 65% of BDs, followed by the *MLL2* gene (mutated in 53% of BDs), with mutations also frequent within the *TP53* and *NOTCH2* genes (in 53% and 47% respectively). Truncating mutations in the *TP53* gene were noted in 4 of 9 BD cases with mutations in this gene.

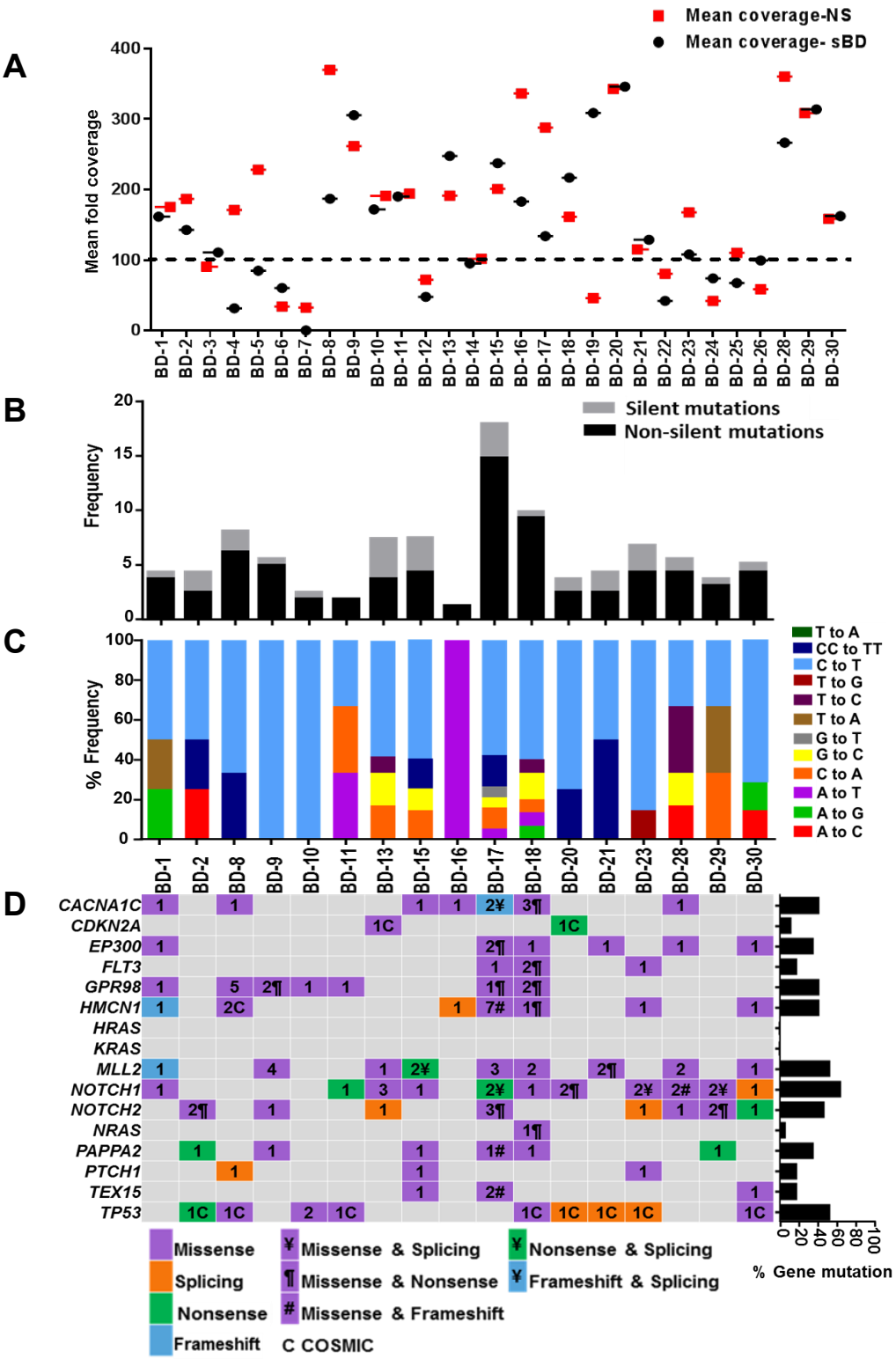


Figure 4.7 Target enriched sequencing data from solitary BD lesions (sBD); part A shows results from the 29 solitary BDs which underwent target enriched sequencing (1-30), whereas parts B – D show results from the 17 sBDs remaining after the second round of filtration. **A:** The mean fold coverage over the 29 BD samples. **B:** Frequency of non-silent: silent (synonymous) mutations (N = 17). **C:** Frequency of different types of base pair alterations (expressed as percentage of overall mutations), showing that C > T is the most common type of base pair change. **D:** Heatmap of all non-silent mutations within the target genes of sBD samples. The numbers within the heatmap indicate the number of mutation/gene/sample with the different colours each indicating a different type of mutation. The mutation frequency of each of the targeted genes is plotted as a corresponding bar graph on the right of heatmap. Note that in B – D, each of the columns represent the same BD sample.

4.4.1.2.3 Data analysis on cSCCs adjacent to potential precancerous lesions

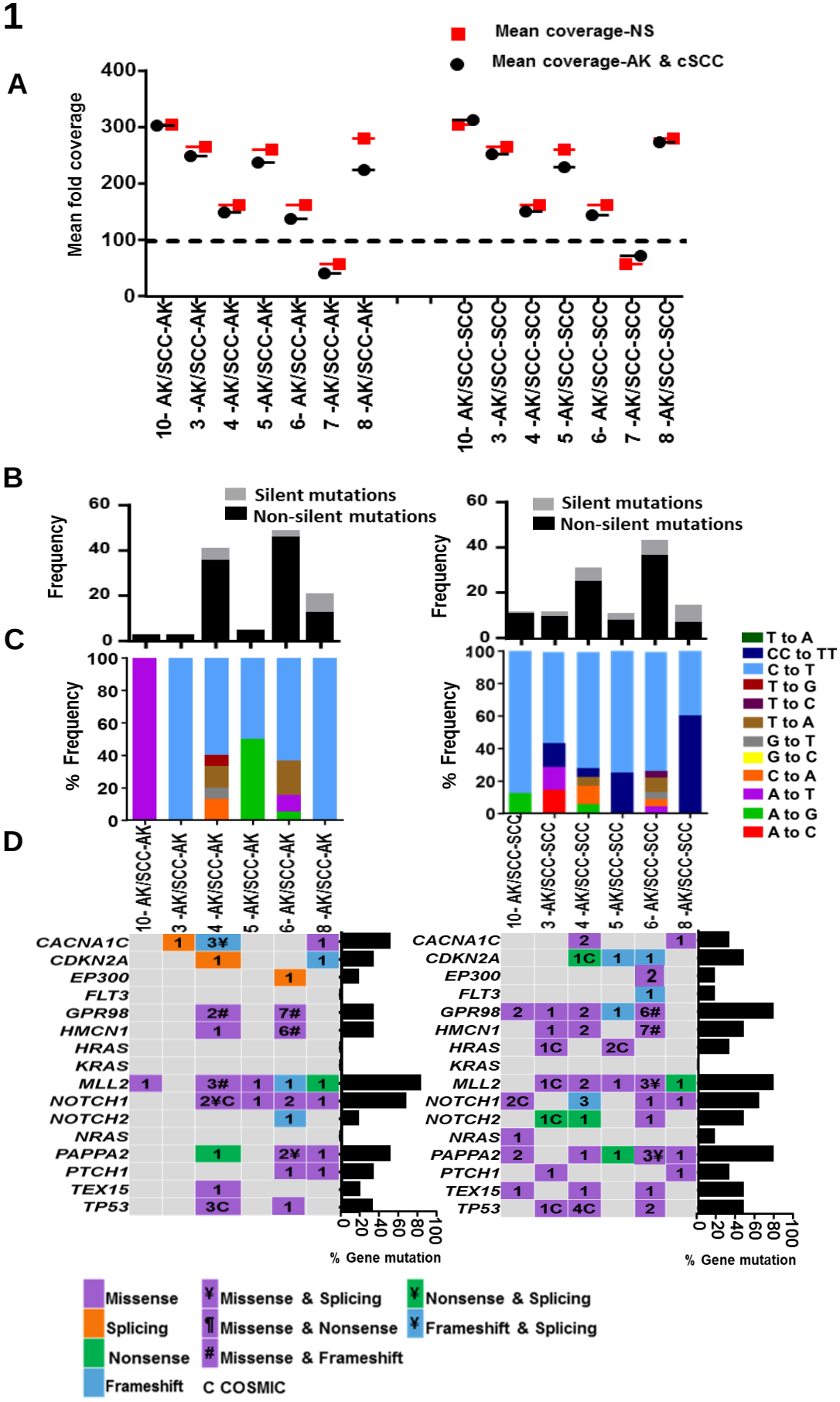
After the second filtration, 6 of 7 AK/cSCCs and 6 of 8 BD/cSCCs remained for analysis. The ratio of non-silent somatic mutations to silent (synonymous) mutations was 3:1 and 2.7:1 within adjacent AK and cSCC lesions respectively and was 3:1 and 2:1 within adjacent BD and cSCC lesions respectively.

A high frequency of C > T mutations was noted within all lesions; median frequency 62% (range 0 – 100 %) for AKs, median frequency 73% (range 40 – 87.5 %) for cSCC adjacent to AK, median frequency 65% (range 34 - 100%) for BDs and 64% (range 50 - 80%) for cSCC adjacent to BDs. Moreover, the median frequency of CC > TT was 0 and 9% for AK and cSCC paired lesions respectively, and 7 and 3% for paired BD and cSCC lesions respectively as shown in figure 4.8.

Consistent with the higher mutation burden within malignancy, cSCCs adjacent to AKs showed a higher total number of mutations than the adjacent AK (73 in cSCCs adjacent to AKs and 49 in AK adjacent to cSCCs), however in the BD/cSCC cases, the BDs had slightly higher mutation burden than the adjacent cSCC (47 and 32 mutations in the BDs and cSCCs respectively) as per figure 4.8.

Overall comparison of the AK/cSCC and BD/cSCC samples showed some variation in the genes which were mutated in paired samples, with additional mutations in the cSCC on occasions, but in other cases an absence of mutations in the cSCC that were present in the AK or BD. For example, three BDs within BD/cSCC lesions showed non-silent somatic mutations (*CDKN2A* in two BDs and *NRAS* in the other BD) that were not seen to be mutated within the adjacent cSCCs. Conversely, in three cSCCs adjacent to AKs, *NRAS* and *HRAS* were mutated only in cSCCs (in one cSCC and two cSCCs respectively) but not in the adjacent AKs. Most of the mutations, as seen on the heat maps in figure 4.8, were missense mutations and to less extent truncating/ nonsense mutations.

1


A
Mean coverage-NS
Mean coverage-AK & cSCC
Mean fold coverage
400
300
200
100
0
10- AK/SCC-AK
3 -AK/SCC-AK
4 -AK/SCC-AK
5 -AK/SCC-AK
6- AK/SCC-AK
7 -AK/SCC-AK
8 -AK/SCC-AK
10- AK/SCC-SCC
3 -AK/SCC-SCC
4 -AK/SCC-SCC
5 -AK/SCC-SCC
6- AK/SCC-SCC
7 -AK/SCC-SCC
8 -AK/SCC-SCC
B
Silent mutations
Non-silent mutations
Frequency
C
% Frequency
T to A
CC to TT
C to T
T to G
T to C
T to A
G to T
G to C
C to A
A to T
A to G
A to C
D
CACNA1C
CDKN2A
EP300
FLT3
GPR98
HMCN1
HRAS
KRAS
MLL2
NOTCH1
NOTCH2
NRAS
PAPPA2
PTCH1
TEX15
TP53
% Gene mutation
Missense
Splicing
Nonsense
Frameshift
Missense & Splicing
Missense & Nonsense
Missense & Frameshift
C COSMIC
Nonsense & Splicing
Frameshift & Splicing

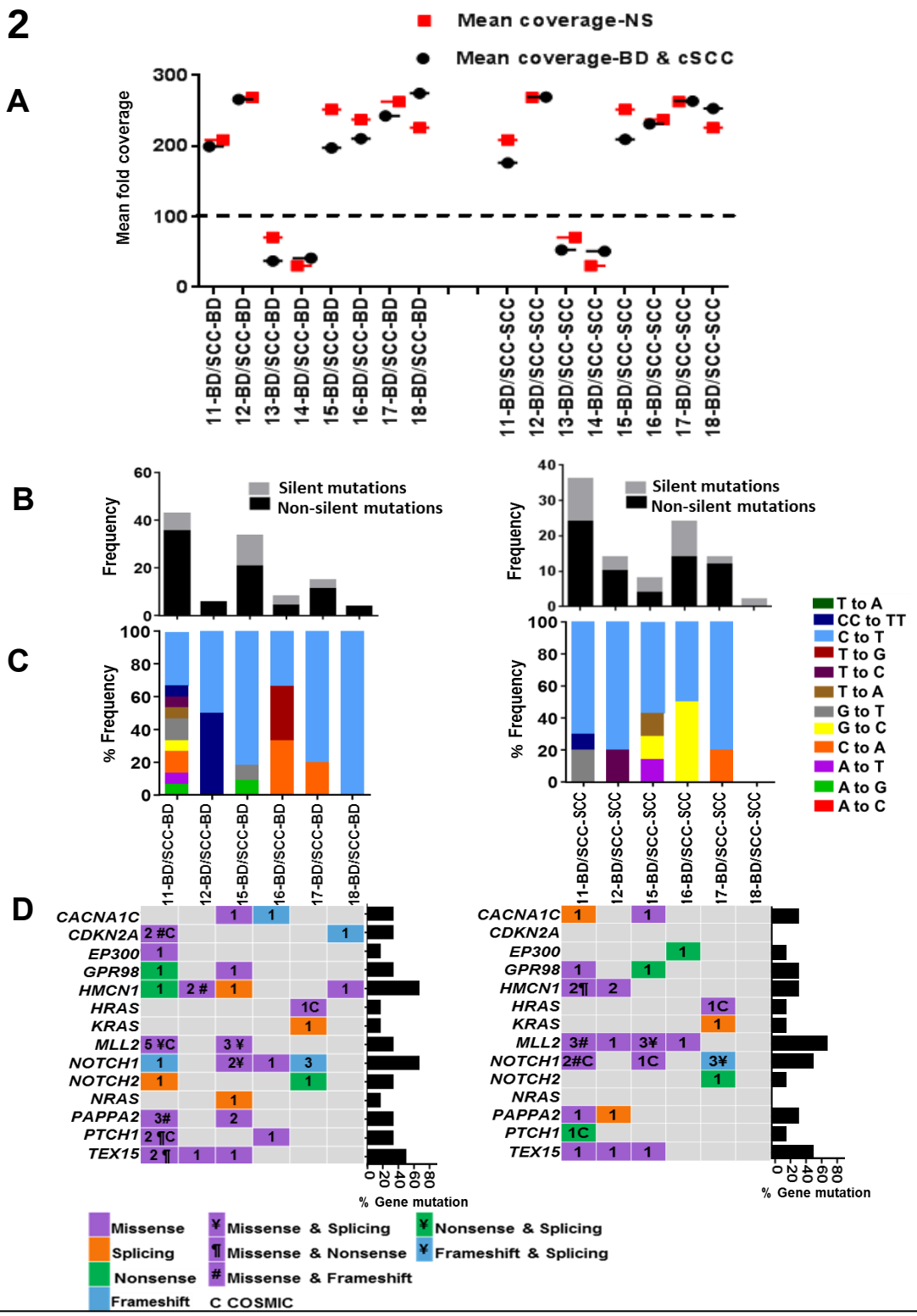


Figure 4.8 Target enriched sequencing results of **1:** Adjacent AK and cSCC lesions. **2:** Adjacent BD and cSCC lesions**.** Parts A in 1 and 2 shows results from the 7 adjacent AK/cSCC) and 8 adjacent BD/cSCC lesions which underwent target enriched sequencing, whereas parts B – D show results from the 6 AK/cSCC and 6 BD/cSCC lesions after the second round of filtration. **A:** The mean fold coverage for the samples. **B:** The frequency of non-silent and silent (synonymous) base pair changes within the lesions. **C:** The frequency of different base pair changes, highlighting C > T alterations as the most common alteration. **D:** Heatmaps highlighting the mutations within the 18 target genes. The numbers within the heatmap indicate the number of mutations/gene for each sample with the different colours indicating a particular type of mutation. The bar graph on the right of the heatmap demonstrates the frequency of mutation for each of these genes as percentages. Note that in B – D, each of the columns represent the same AK, BD or cSCC sample. AK/SCC-AK = AK part of the paired lesion. AK/SCC-SCC= SCC part of the paired lesion. BD/SCC-BD = BD part of the paired lesion. BD/SCC-SCC = SCC part of the paired lesion.

The number of non-silent mutations within the target sequencing study is summarised in table 4.5.

Table 4.5 Non-silent somatic mutation results from the target enriched sequencing data following filtration. The number of samples and the number of non-silent mutations are presented for each of the various groups (solitary AK, solitary BD and adjacent lesions (cSCC and BD/AK).

	Solitary AK	Solitary BD	AK adjacent to cSCC	BD adjacent to cSCC	cSCC adjacent to BD	cSCC adjacent to AK
Number of samples	18	17	6	6	6	6
Total number of non-silent mutation	308	122	49	47	32	73

The deleterious effect of the non-silent mutations for each gene was predicted using (SIFT)(www.sift.bii.a-star.edu.sg) and the results are summarised in table 4.6. Many of the genes showed high number of non-silent mutations with deleterious effect (the highest is within *MLL2* and *GPR98* while most of the mutations on *PTCH1* gene were tolerable).

Table 4.6 The proportion of deleterious mutations within targeted genes per sample. The predicted effect of the mutations was according to the result of SIFT and the proportion of the deleterious mutations is plotted per the total number of non-silent mutation per gene. The (-) sign indicates either no detected mutation or the detected mutations were not deleterious within the targeted genes.

Gene ID	Solitary AK	Solitary BD	AK adjacent to cSCC	cSCC adjacent to AK	BD adjacent to cSCC	cSCC adjacent to BD
CACNA1C	12/30	5/10	2 /5	2/3	0/2	1/2
CDKN2A	2/10	2/2	0 /2	0/3	0/3	-
EP300	1/9	1/7	0/1	0/2	1/1	0/1
FLT3	5/11	0 /4	-	0/1	-	-
GPR98	23/52	10/13	1/9	4/12	2/2	1/2
HMCN1	8 /32	4 /14	1/7	2/10	1/5	0/4
HRAS	2 /3	-	-	3/3	1/1	1/1
MLL2	28/49	12/18	4/7	5/8	6/8	6/8
KRAS	1/1	-	-	-	0/1	0/1
NOTCH1	7/38	7/18	1/6	3/7	1/7	0/6
NOTCH2	4/17	3 /12	0/1	0/3	1/2	1/1
NRAS	-	1/1	-	1/1	0/1	-
PAPPA2	12/19	2/ 6	2/4	3/8	2/5	0/2
PTCH1	1/11	0/3	0/2	0/2	0/3	0/1
TEX15	12 /14	3 / 4	1 /1	2 / 3	4/ 4	3 / 3
TP53	4 /12	5 / 10	1 / 4	3 / 7	-	-

Most of the genes that were mutated in the solitary AK and BD lesions were also mutated in the AKs and BDs adjacent to cSCCs. By combining lesions that shared the same

diagnosis (e.g. solitary AKs and AKs adjacent to cSCC, etc.), one can see that the vast majority of genes that were mutated in the AK and BD lesions were also mutated in cSCCs in this study (table 4.7); this is not surprising because the 18 target genes had been originally selected because they had been reported as mutated in cSCC previously and/or were mutated in the WES results of AKs as documented in chapter 3.

Table 4.7 Percentage of number of samples with mutations in each of the 18 targeted genes within all lesional groups. The groups were as follows, all AKs combined (i.e. solitary AKs and AKs adjacent to cSCC), all BDs combined (i.e. solitary BDs and BDs adjacent to cSCC) and cSCCs (adjacent to AKs or BDs).

	CACNA1C	*CDKN2A*	*EP300*	*FLT3*	*GPR98*	*HMCN1*	*HRAS*	*KRAS*	*MLL2*	*NOTCH1*	*NOTCH2*	*NRAS*	*PAPPA2*	*PTCH1*	*TEX15*	*TP53*	*FAT1*	*DNAH5*
AK	71	33	21	29	58	63	13	4	79	71	50	0	54	29	46	42	0	0
BD	39	22	30	13	39	48	4	4	48	65	43	9	35	17	26	39	0	0
cSCC	33	25	17	8	58	42	25	8	75	58	33	8	58	25	50	25	0	0

4.4.1.2.4 Clonal evolution of the detected mutations

In studies of somatic alterations in neoplastic and pre-neoplastic lesions, it is considered that variants with high allele frequency (i.e. clonal) are likely to have occurred early (Parry et al., 2015). Therefore clonality assessment was performed on the samples within this study. The measurement for clonal evolution assessment was adapted from Parry et al. (2015) and was done under supervision of Dr Matthew Rose-Zerilli, Cancer Research Group, Faculty of Medicine, University of Southampton.

Clonal mutations, defined as the variant allele frequency (VAF) adjusted for tumour purity equal to or larger than 0.45, were determined by a binomial distribution using the number of alternative allele reads, the total number of reads from lesional cells (approximately, total number of reads × purity estimate), and the expected VAF of 0.45. Mutations with cumulative probability distribution function $p < 0.07$ were identified as subclonal (the threshold of $p < 0.07$ was selected according to previous work in the University of Southampton by Parry et al. (2015), which was originally based on the ABSOLUTE method reported by Carter et al. (2012)). As summarised in figure 4.9, there was no obvious pattern of clonal mutations in any of the lesion groups.

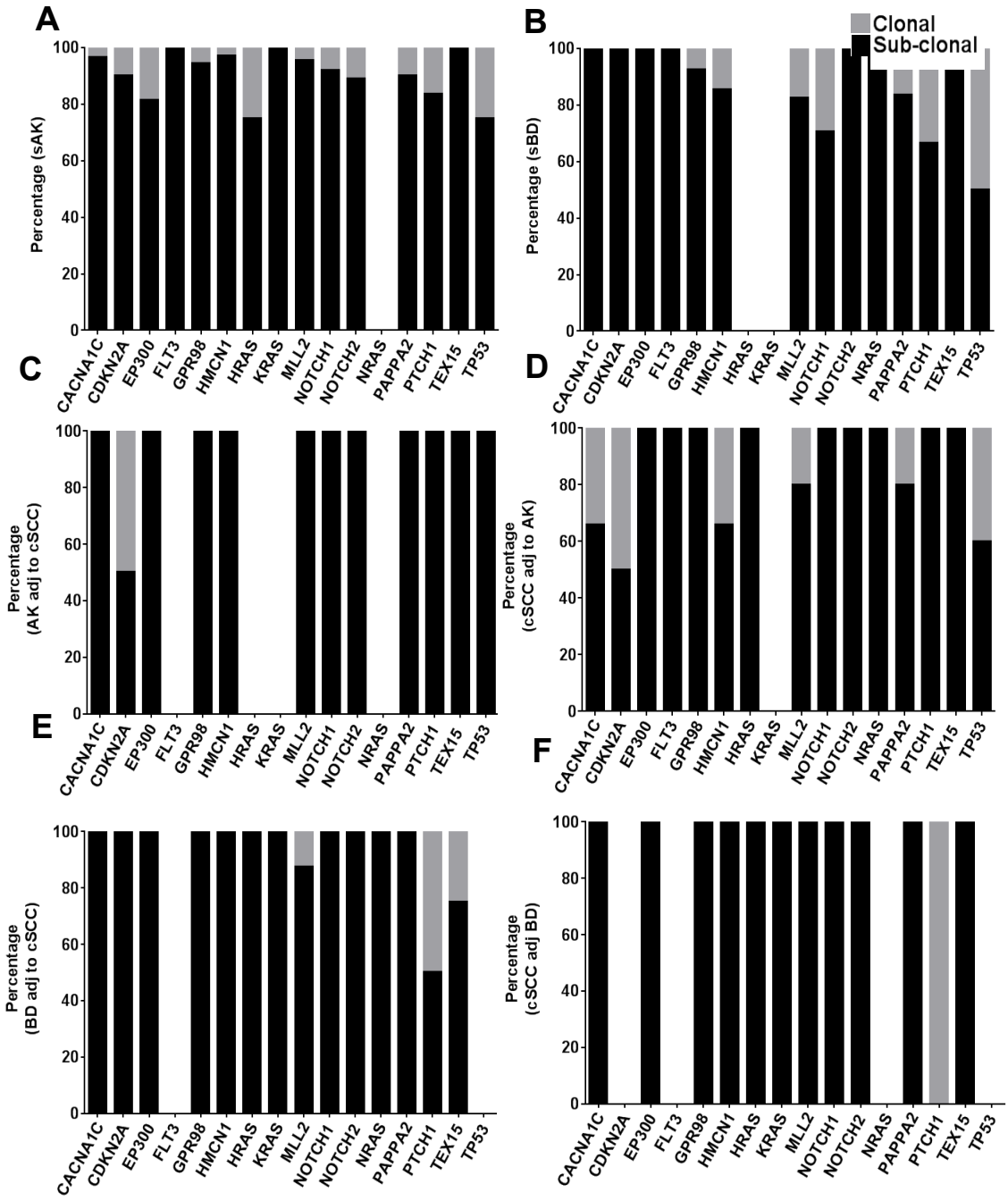


Figure 4.9 Clonality analyses of gene mutations. **A:** solitary AK, **B:** solitary BD, **C:** AK adjacent to cSCC, **D:** cSCC adjacent to AK, **E:** BD adjacent to cSCC, and **F:** cSCC adjacent to BD. Distribution of the estimated proportion of cells harboring a mutation in different sample groups, based on the availability of purity information. Genes are displayed from left to right in alphabetical order. For each gene, the bar graph shows the percentage of clonal/subclonal mutations by using a binomial distribution based on the alternative allele read count, the total read count from lesional cells and an expected variant allelic fraction (VAF) of 0.45. AK = actinic keratosis. BD = Bowen's disease. cSCC = cutaneous squamous cell carcinoma. Adj = adjacent.

Overall, clonal mutations were detected within 13 of the 18 selected genes (*CACNA1C, CDKN2A, EP300, GPR98, HMCN1, HRAS, MLL2, NOTCH1, NOTCH2, PAPPA2, PTCH1, TEX15* and *TP53*) with *TP53* showing the highest number of clonal mutations (12 clonal mutations) within all lesions together.

4.4.2 *KNSTRN* gene Sanger sequencing analysis

The first exon of *KNSTRN* (317 base pairs in length) was amplified by PCR and sequenced using Sanger sequencing on the solitary AKs, solitary BDs and cSCC adjacent to either AK or BD (N = 69) and the matched normal skin. The results were analysed using SeqBuilder (v10.0.1.3) and Seq Scanner (v2.0) software (Applied Biosystems). Only bases with a quality score (Phred score) ≥ 20 were included in the analysis. Out of the 69 samples, the initial analysis suggested that 5 of 25 AKs, 4 of 29 BDs and 0 of 7 cSCCs adjacent to AKs, and 1 of 8 cSCCs adjacent to BD contained a p.Ala40Glu mutation in *KNSTRN* exon 1. According to the Variant Effect Predictor tool on the Ensembl webpage (Ensembl)(www.ensembl.org/info/docs/tools/vep/) this alteration has been reported previously as a missense variant (c.119C > A within codon GCG) and according to the functional prediction programmes (SIFT and PolyPhen2) this variant is expected to be deleterious. As a confirmatory step, the reverse strands of DNA from the samples with this alteration were also sequenced and mutations were confirmed in the reverse strand in all cases. In most of the cases the base pair change was heterozygous where the alteration occurred on one allele only, however homozygous alteration was also seen within 2 of the 5 AKs and 1 of the 4 BDs with *KNSTRN* base alterations (figure 4.10).

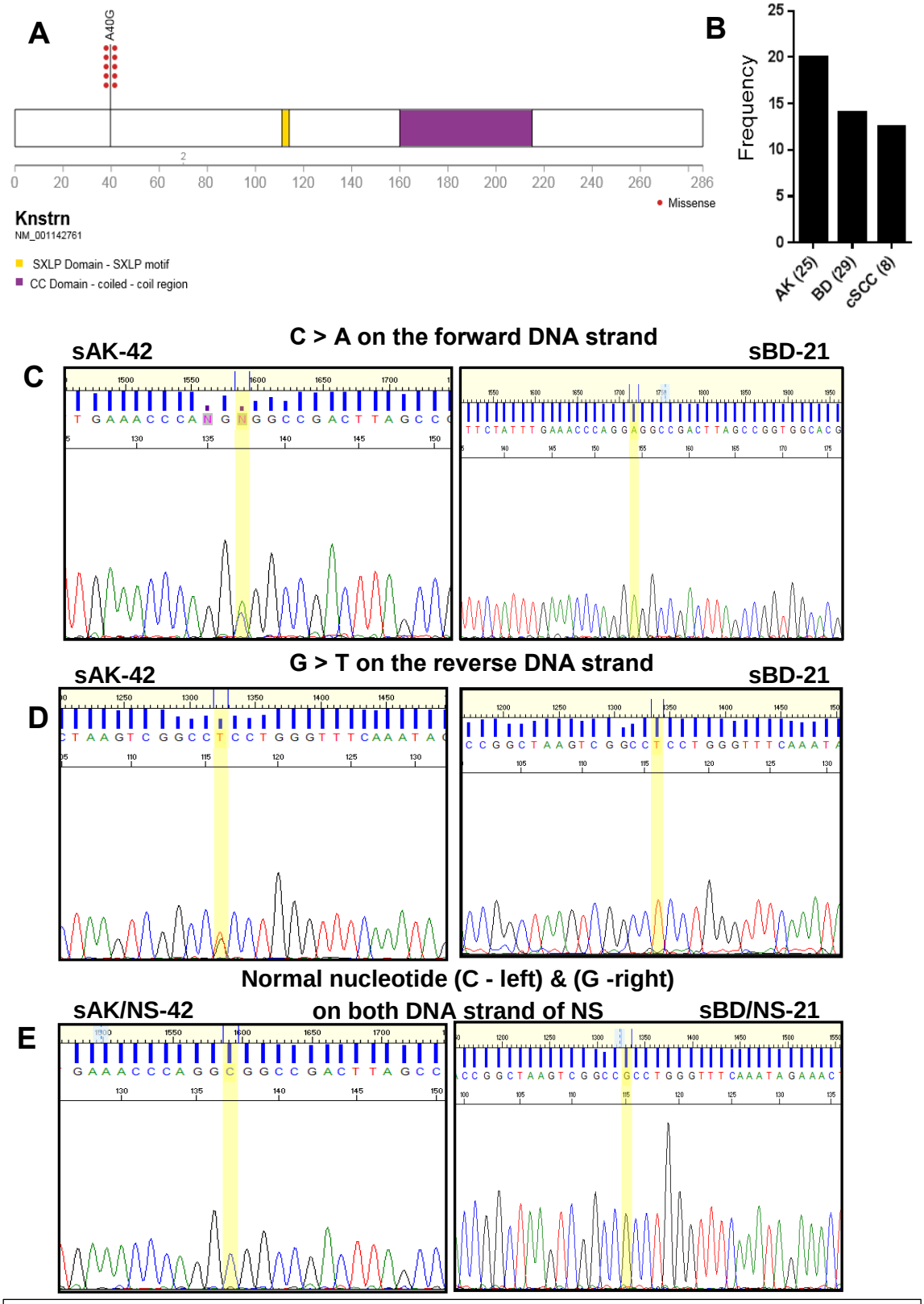


Figure 4.10 *KNSTRN* gene mutations. **A:** Position of the p.Ala40Glu mutation in the *KNSTRN* coding sequence. **B:** Frequency of *KNSTRN* mutation encoding Ala40Glu in actinic keratosis (AKs), Bowen's disease (BDs) and cSCC adjacent to BD. **C – E:** Sequencing analysis using SeqScanner (v2.0) software showing; **C:** C > A on the forward DNA strand where the change is heterozygous on the left side and homozygous on the right side. **D:** The results on the reverse DNA strand (G > T) are shown as heterozygous on the left and homozygous on the right. **E:** Normal genotype on the matched normal skin (forward strand on the left side and reverse strand on the right side) is also shown. NS = Normal skin. Nucleotide colour; A=Green. T=Red. C=Blue. G=Black.

4.5 Discussion

At the later stages of skin carcinogenesis, the evidence suggests that a number of cSCCs arise from AKs and BD (Cox et al., 1999, Criscione et al., 2009, Kao, 1986, Marks et al., 1988, Salasche, 2000). Additional support for this includes the similarity in the pattern of expression of different proteins, including p53 and beta catenin, between cSCCs, BD and AKs (Brasanac et al., 2005, Einspahr et al., 1997, Papadavid et al., 2002). Moreover, single gene base pair analysis experiments have shown similar genetic alterations within *TP53, CDKN2A* and *RAS* family genes between cSCCs and AKs (Kanellou et al., 2008, Kubo et al., 1994, Lee et al., 2000, Oberholzer et al., 2012, Ziegler et al., 1994). The results of the WES study from chapter 3 revealed a similarity between cSCCs in the published literature and the AKs in this study in the frequency and burden of mutated genes (Durinck et al., 2011, Li et al., 2015, Pickering et al., 2014, South et al., 2014), but the WES analysis was limited to 5 AK samples. Thus to overcome this limitation, targeted sequencing on a list of 18 genes at higher sequencing depth (x = 200) was conducted on wider range of samples that included AKs and BDs, as well as a smaller group of cSCCs. Up to the date of conducting this analysis, this was the first study to conduct target enriched sequencing on AKs and BDs.

In addition to confirming the presence of mutations in *TP53, CDKN2A* and *RAS* family genes in AKs and *TP53* in BDs (Campbell et al., 1993a, Campbell et al., 1993b, Kanellou et al., 2008, Nelson et al., 1994, Nindl et al., 2007, Pacifico et al., 2008, Park et al., 1996, Soufir et al., 1999, Taguchi et al., 1998, Takata et al., 1997, Zaravinos et al., 2010, Ziegler et al., 1994), AKs and BDs showed mutations within most of the other targeted genes, with most of the mutations scattered over all exons of the relevant genes.

From the data, precancerous lesions (AK and BD) showed mutations in most of genes that were also seen to be mutated in cSCC in this study. As *NOTCH1, NOTCH2* and *TP53* have been discussed earlier in this thesis, the following discussion is focussed mainly on the other genes which were frequently mutated in the AKs and/or BDs in this chapter.

The *CACNA1C* gene was identified by South et al. (2014) amongst the genes with a high mutation frequency within cSCC (as were *NOTCH1, NOTCH2* and *TP53*) and was also reported as mutated within aggressive cSCCs (Pickering et al., 2014). The *CACNA1C* gene encodes an alpha-1 subunit of a voltage-dependent calcium channel and mediates the

influx of calcium ions into the cell upon membrane polarisation (NCBI) (www.ncbi.nlm.nih.gov/gene). In addition to cSCC, mutations of the *CACNA1C* gene have been reported within cutaneous melanoma (Berger et al., 2012) and *CACNA1C* gene base pair alterations is associated with shorter QT interval on electrocardiogram (as a result of its expression in cardiac muscle) (Antzelevitch et al., 2007).

The *HMCN1* gene was also mutated frequently in the AKs, BDs and cSCCs. *HMCN1* has been reported to be mutated in cSCC (Durinck et al., 2011) and melanoma (Nikolaev et al., 2012). As a result of its role in hemidesmosomes arrangement within the epidermis, *HMCN1* gene mutations have been considered to have biological significant in the disease pathology of these cancers, with loss of function mutations possibly facilitating metastasis through weakening adhesion of the cells to the basement membrane (Feitosa et al., 2012). However, it is unclear at this stage how the presence of *HMCN1* mutations in AKs and BDs affects their behaviour.

In the skin, genetic alterations within chromatin remodelling genes (*EP300* and *MLL2*) are also seen in the target enriched sequenced samples. *EP300* is a histone acetyltransferase gene and is known to act as a transcriptional coactivator of *NOTCH* pathway genes. *MLL2* is a histone methyltransferase gene and which methylate the Lys-4 position of histone H3 and transcriptionally regulates beta-globin and oestrogen receptor genes (NCBI) (www.ncbi.nlm.nih.gov/gene). *MLL2* mutation has been detected in malignant melanoma (Berger et al., 2012) as well as in aggressive cSCCs (Pickering et al., 2014), and is frequently mutated in non-Hodgkin lymphomas (Morin et al., 2011). *PAPPA2* encodes a member of the pappalysin family of metalloproteinases and is known to play a role in regulation of insulin-like growth factor (NCBI) (www.ncbi.nlm.nih.gov/gene). *PAPPA2* has been reported to be mutated in cSCC (Durinck et al., 2011) and head and neck SCC (i.e. oropharyngeal SCC) (Stransky et al., 2011). *GPR98* is a member of the G protein-coupled receptor family which have 7 transmembrane domains that allow cells to respond to extracellular signals. G protein-coupled receptor proteins, including *GPR98*, are expressed widely in the central nervous system (CNS) and mutations on that gene family is associated with hearing loss, visual impairment and with seizures. Within skin, *GPR98* mutations have been detected in malignant melanoma (Prickett et al., 2011) and in aggressive cSCCs (Pickering et al., 2014).

TEX15 mutations were reported previously in malignant melanomas (Berger et al., 2012, Nikolaev et al., 2012) and within cSCCs (Durinck et al., 2011, Pickering et al., 2014). In addition, mutations of *TEX15* have been associated with oligozoospermia and/or azoospermia in Caucasian men (Aston et al., 2010). *PTCH1* gene mutations are well known to be associated with the development of BCCs (Evans et al., 2000, Reifenberger et al., 2005, Reifenberger et al., 1998, Zhang et al., 2001a), however, *PTCH1* mutation were detected previously within cSCCs (Durinck et al., 2011, Ping et al., 2001).

Whereas most of these genes have been reported as mutated in cSCC, additional supporting evidence for a role of many of these genes in skin biology comes from the fact that most of these genes are known to be expressed at the mRNA and/or protein levels in skin, as summarised in table 4.8 (The Human Protein Atlas) (www.proteinatlas.org/), (Berglund et al., 2008, Uhlén et al., 2005).

Table 4.8 Protein and mRNA expression for the 18 selected genes within skin tissue. This information was obtained from The Human Protein Atlas website and the results are based on two published studies (Berglund et al., 2008, Uhlén et al., 2005). N/A = Not applicable (i.e. no result for this gene has been published yet). mRNA = messenger DNA.

Gene ID	Protein expression	mRNA expression
CACNA1C	YES	YES
CDKN2A	YES	YES
DNAH5	NO	NO
EP300	YES	YES
FAT1	YES	YES
FLT3	YES	YES
GPR98	N/A	N/A
HMCN1	YES	YES
HRAS	YES	YES
KRAS	YES	YES
MLL2	YES	YES
NOTCH1	YES	YES
NOTCH2	N/A	YES
NRAS	YES	YES
PAPPA2	NO	NO
PTCH1	YES	YES
TEX15	NO	NO
TP53	YES	YES

Clonality assessment within this chapter demonstrated clonal mutations in 13 of the 18 selected genes, suggesting that the mutation in the relevant gene could have been occurred early during the development of the AK or BD. However, it was difficult to recognise any specific pattern of clonality within each of the lesions, thus it seemed that mutations in different genes could account for one of the early genetic events in the

development of these precancerous lesions and that there was not a single sequential genetic mutation pathway which led to the growth of most of these lesions.

The *KNSTRN* p.Ala40Glu mutation was detected in 20%, 14% and 12.5% of AKs, BD, and cSCCs respectively. According to Lee et al. (2014) the p.Ser24Phe mutation was the most frequent mutation noted in 5 of 27 (19%) AKs and 5 of 38 (13%) cSCCs in their study as well as in 44% of cutaneous melanomas (Lee et al., 2014). None of the samples investigated in this chapter showed any p.Ser24Phe mutation and the only mutation that was seen was the Ala40Glu, which was reported also in cSCCs in the Lee et al. (2014) publication but at lower frequency than p.Ser24Phe (in that paper the exact frequency of the Ala40Glu change was not shown). Functional analysis on the effect of the detected mutations on the function of the protein was conducted by Lee et al. (2014) and chromosome segregation during mitosis was disrupted within primary human keratinocytes expressing mutant *KNSRTN* gene, with the percentage of unpaired chromatids significantly higher in the p.Ala40Glu transduced primary human keratinocytes (Lee et al., 2014).

Overall, in this study, precancerous AK and BD contained a high frequency of mutations within the selected genes, with clonality analysis suggesting that some of these mutations may have occurred at an early stage during development of these lesions. Thus, investigating for mutations in these genes in chronically sun exposed normal looking skin would help to determine whether this was indeed the case (as detailed in the next chapter).

Chapter 5: Target enriched sequencing on p53 immunopositive patches in human epidermis

5.1 Introduction

The results of target enriched sequencing on the precancerous skin lesions, AK and BD, in the previous chapter (chapter 4) of this thesis demonstrated a moderate to high frequency of mutations in the 18 selected genes in these lesions. In addition to these clinically visible precancerous skin lesions, it is known that chronically sun exposed skin contains earlier potentially precancerous lesions consisting of clinically invisible patches of epidermal keratinocytes that stain positive for p53 protein (Berg et al., 1996, Jonason et al., 1996, Ponten et al., 1995, Rebel et al., 2001, Ren et al., 1996). These "p53 immunopositive patches" (PIPs) are seen in chronically UV irradiated skin in humans and mice, with DNA sequencing identifying *TP53* mutations in 29-64% of these patches (Jonason et al., 1996, Kramata et al., 2005, Robinson et al., 2010). As both cSCCs and PIPs share UVR as the main risk factor and as similar *TP53* gene mutations have been reported in both PIPs and cSCCs (Berg et al., 1996), it has been hypothesised that "PIPs" may be an early precursor of cSCC (Brash et al., 1991, Kramata et al., 2005, Rebel et al., 2005).

PIPs have been defined as un-interrupted clusters of ≥ 10 keratinocytes with homogenous and strong immunoreactivity to anti-p53 protein antibody that are distributed within the basal and suprabasal epidermis (De Graaf et al., 2008, Ren et al., 1997). However, Zhang et al. (2001b) included PIPs with as few as three p53 positive epidermal cells in a study in which the number and size of PIPs were positively correlated with the dose and duration of UVR exposure (Zhang et al., 2001b). Related to this, there is evidence that the incidence of PIPs increases with age and within skin adjacent to cSCCs (Rebel et al., 2005).

Previous studies of PIPs have studied the genetic changes within *TP53* (Brash et al., 1991, Tabata et al., 1999, Ziegler et al., 1994). To date, there is no data on somatic DNA mutations within other genes in PIPs. Cancer in general is a multistep process, usually starting with one initial mutation that expands clonally (Foulds, 1954, Loeb, 1991, Vogelstein et al., 2013). A recent study by South et al. (2014) showed *NOTCH1* and *NOTCH2* gene mutations within normal looking skin adjacent to cSCC lesions, but they did not identify PIPs within these skin samples, so it is not known whether these mutations occurred in PIPs, or within the non-PIP epidermis.

Thus, to extend the knowledge on cSCC pathogenesis, and to determine whether PIPs contain mutations in genes relevant to development of AKs, BDs and cSCCs, target

enriched sequencing of the 18 genes that were sequenced in chapter 4 to be conducted on PIPs.

5.2 Hypothesis and aims

The hypothesis was that PIPs may contain mutations in genes that play a role in the development of AKs, BDs and/or cSCCs. Therefore the aim of this chapter was to define genetic alterations within PIPs through performing target enriched sequencing on the 18 genes panel which was sequenced in AKs and BDs in chapter 4.

5.3 Materials and Methods

5.3.1 Tissue sample

Normal looking skin which was at least 4 mm away from cSCC or BCC lesions were obtained from Dr. Chester Lai, Dermatology Department, University Hospital Southampton NHS Foundation Trust. As described in chapter 3, blood DNA samples were obtained from patients as part of an investigation on skin cancer being conducted by Dr Lai. The study was conducted under local research ethics committee approval and a signed consent form was obtained from all subjects for the use of the tissue.

5.3.2 Epidermal sheet separation and immunostaining

Epidermal sheets were separated from the underlying dermis as discussed in the materials and methods (section 2.2.3). Epidermal sheets were then stained using DO-7 anti-p53 antibody as described in materials and methods (section 2.2.4). The aim of immunostaining was to define the PIPs within the whole mount epidermis. The extent of the p53 staining within the epidermis was also examined within vertical sections (5µm thick) through the PIPs to determine the proportion of cells staining for p53 protein.

5.3.3 Haematoxylin & Eosin (H&E) staining

H&E staining was performed, as described in section 2.2.5 on 5 µm thick vertical sections of epidermal sheets that were positive for PIPs. The main purpose for this H&E counterstain was to allow the determination of the numbers / proportion of epidermal nuclei which were immunopositive for p53.

5.3.4 PIP microdissection

PIPs were dissected as described in section 2.2.8. Dissected PIPs and the surrounding p53 negative epidermis were collected separately within microcentrifuge tubes containing 180 µl of lysis (ALT) buffer from the QIAamp DNA FFPE Tissue kit (Qiagen).

5.3.5 DNA extraction

Genomic DNA was extracted from the PIPs and matched p53-immunonegative normal looking epidermis which had been microdissected as above. As outlined in section 2.2.9, genomic DNA was purified using a QIAamp DNA FFPE Tissue Kit (Qiagen) according to the manufacturer's instructions.

5.3.6 Assessment of DNA concentration within the samples

A Qubit 2.0 fluorometer (Invitrogen) was used for quantification of genomic DNA, as described in section 2.2.10.2, on DNA isolated from the PIPs and non-PIPs epidermis. Target DNA concentration for the target enriched sequencing analysis was ≥ 5 ng/µl, with a target total DNA amount of ≥ 100 ng.

5.3.7 Assessment of DNA quality

Using Nanodrop D-1000 (Nanodrop Technologies, US) absorbance spectrophotometer, the optical density in the form of A260/A280 and A260/A230 ratios was measured for each DNA sample as detailed in section 2.2.11. DNA fragmentation was assessed as detailed in section 2.2.11.1 using a panel (multiplex) PCR that amplifies different DNA fragment lengths (100, 200, 300, 400 and 600 bps). The resultant bands were visualised on an agarose gel electrophoresis as detailed in section 2.2.11.2.

5.3.8 Target enriched sequencing

Target enriched sequencing was performed as described in chapter 4 (section 4.3.7).

5.3.9 *KNSTRN* PCR and Sanger sequencing

KNSTRN PCR and Sanger sequencing was conducted as described in section 4.3.8.

5.3.10 Statistics

GraphPad Prism (v6.0) was used to generate figures and to undertake statistical analysis (support was received from Scott Harris and Ho-Ming (Brian) Yuen, Medical Statistics, Primary Care and Population Sciences, Faculty of Medicine, University of Southampton).

5.4 Results

Whole mount epidermis from fresh normal looking skin samples at least 4 mm away from cSCC or BCC lesions, from 126 subjects, were stained for p53. Information on the age, gender, number of subjects with PIPs in the epidermis sample (and range of PIPs per subject), number of PIPs that underwent DNA extraction, DNA concentration and amount obtained from samples where DNA was extracted, PIP samples of sufficient size that gave the targeted DNA concentration, number of PIPs that fulfilled the quality and quantity criteria and underwent target enriched sequencing, number of PIPs that gave primary sequencing results and number of PIPs that pass the second filtration criteria (as described for secondary filtration of AK and BD data in sections 4.3.7.3 & 4.4.1.2) are provided in table 5.1 and 5.2 (later in this chapter). The relevant information for each of the steps which provided the data in this table is presented in the subsequent paragraphs.

Table 5.1 Identification and selection of PIPs for target enriched sequencing.

Total number of subjects	126
Age	Range 44 - 95 years, Median 83 , Mode 82
Gender	F = 55 M = 71
Number of subjects with PIPs in their epidermis sample	89 patients
Range of PIPs	Range 1 – 8 distributed as the following
	40 subjects have 1 PIPs each
	20 subjects have 2 PIPs each
	5 subjects have 3 PIPs each
	11 subjects have 4 PIPs each
	6 subjects have 5 PIPs each
	3 subjects have 6 PIPs each
	2 subjects have 7 PIPs each
	2 subjects have 8 PIPs each
Mean range of PIPs per subject	217/89 = 2.4 PIPs per subject
Number of PIPs that underwent DNA extraction (PIPs size range between 0.1 – 0.27 mm² and estimated to have ≥ 1500 p53 immunopositive keratinocytes)	40

The anti-p53 antibody (DO-7) (Dako/UK) was used to detect nuclear p53 protein staining within wholemount epidermal sheets (Guinea-Viniegra et al., 2012). PIPs were defined according to Ren et al. (1997) and De Graaf et al. (2008) (at least 10 homogenous and strongly p53 positive nuclei within a cluster and very clearly separated from the

surrounding epidermis). Epidermal sheets that showed a lot of background staining were excluded. Examples of p53 staining are shown in figure 5.1. The quality of the staining was assessed by 3 observers (Prof Eugene Healy, Dr Chester Lai and myself) in order to have consensus in designating positively stained areas as PIPs.

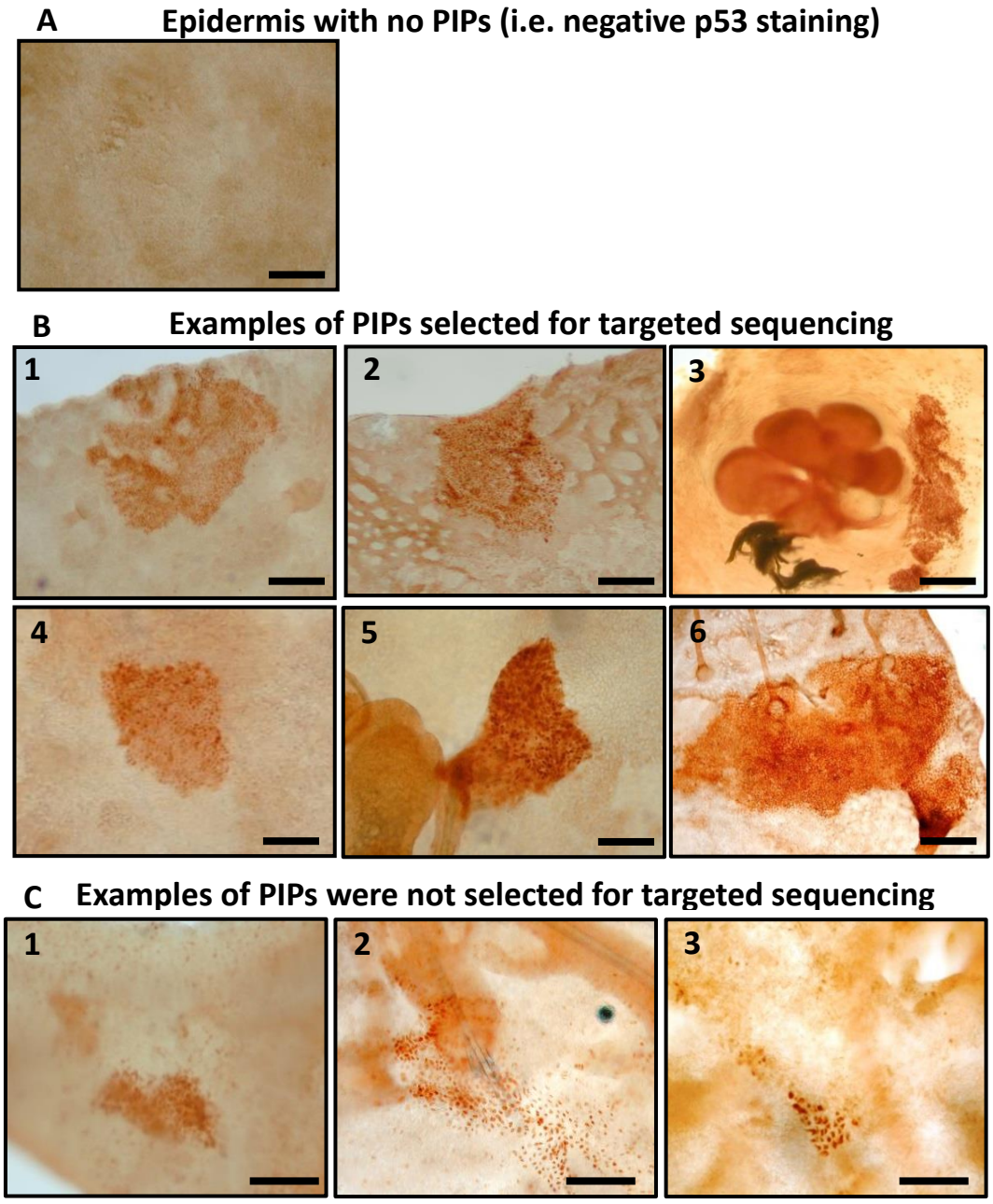


Figure 5.1 p53 staining of epidermal sheets. **A:** An example of an epidermis without any PIPs (i.e. negative p53 staining). **B:** 6 different examples of PIPs (1 – 6) that were selected for target enriched sequencing. **C:** 3 different examples of PIPs (1 -3) that were not selected for target enriched sequencing (due to their small size and the fact that the margin was not completely well demarcated from the adjacent epidermis). The view in each case is from the basal surface of the epidermis. 4x represents fold magnification of images. Scale bar; 200 µm.

The approximate numbers of p53 positive keratinocytes within each of the PIPs were counted using 40 fold magnification under a Nikon Eclipse 80i light microscope (as detailed in section 2.2.6). The cells within the patches were counted from the basal surface of the epidermal sheet and the S10 stage counting grid was used to help in the counting process and reduce the chance of cell recounting. The cells within each PIP were counted 3 times and the mean was calculated. Moreover, to check the accuracy of the results, a second observer counted several patches separately and a comparison of the results suggested that the mean value for the cells counted within the selected PIPs were accurate. ImageJ software analysis using 4x images from each of the PIPs was used for measuring the area of the patches as detailed in sections 2.2.6 and 2.2.7. In cases where the patch size seemed to be small and unlikely to give sufficient DNA, the patches were omitted and were not included for any further analysis; this was the situation for the majority of the patches. On the other hand, where the patches were deemed large enough to potentially provide sufficient DNA, the size of the patches were measured and ranged from 0.10 – 0.27 mm^2 (median 0.145 mm^2) and the number of cells within the patches ranged from 1936 – 4751 cells (median 2597 cells/PIP). Out of the 40 PIPs that underwent DNA extraction, 23 PIPs (identified from 15 subjects) gave sufficient yield of DNA (total DNA amount of ≥ 100 ng) and were selected for target enriched sequencing (table 5.2).

Table 5.2 PIPs underwent DNA extraction. Selected PIPs for target enriched sequencing were plotted in the first part of the table. The area of each clone was measured using ImageJ software and the numbers of cells within each PIP counted under light microscopy.

PIP ID	Area of PIP in mm²	Number of cells per PIP	DNA concentration in ng/µl	Total DNA amount in ng within 15µl	Selected PIPs for target enriched sequencing
1	0.14	2446	6.8	102	Yes
2	0.14	2479	7.1	106.5	Yes
3	0.14	2592	7.43	111.4	Yes
4	0.15	2602	8.89	133.3	Yes
5	0.15	2732	9.1	136.5	Yes
6	0.16	2952	10.1	151.5	Yes
7	0.16	2948	10.3	154.5	Yes
8	0.17	2963	11	165	Yes
9	0.17	2955	12	180	Yes
19	0.18	2965	15.7	235.5	Yes
11	0.18	2977	16.1	241.5	Yes
12	0.18	2983	16.3	244.5	Yes
13	0.19	2703	16.5	247.5	Yes
14	0.19	2993	18.6	279	Yes
15	0.19	2756	19.8	297	Yes
16	0.21	2721	20	300	Yes
17	0.21	2892	23.4	351	Yes
18	0.22	3348	25	375	Yes
19	0.23	3479	29	435	Yes
20	0.24	4572	31.3	469.5	Yes
21	0.24	4682	36.9	553.5	Yes
22	0.25	4532	37	555	Yes
23	0.27	4751	37.3	559.5	Yes
24	0.1	1936	4.40	66	No
25	0.1	1973	4.87	73	No
26	0.1	1989	5.20	78	No
27	0.1	1994	5.33	80	No
28	0.1	1995	5.40	81	No
29	0.1	2001	5.47	82	No
30	0.11	2004	5.60	84	No
31	0.11	2015	5.67	85	No
32	0.11	2034	5.80	87	No
33	0.11	2052	5.87	88	No
34	0.11	2061	5.93	89	No
35	0.11	2077	5.93	89	No
36	0.11	2089	6.20	93	No
37	0.11	2109	6.33	95	No
38	0.12	2201	6.53	98	No
39	0.12	2231	6.60	99	No
40	0.13	2298	6.61	99.1	No

To assess whether the p53 staining within the PIPs stained all of the nuclei within that part of the skin (i.e. as a way of estimating the likely purity of the p53-positive cells within the PIPs), 20 epidermal samples with PIPs were frozen in OCT and 5 µm thick vertical sections were cut through the PIPs with a cryostat microtome. The vertical extent of the p53 staining was assessed within these PIPs and a number of these p53 positive vertical sections were stained with H&E for counter staining (figure 5.2).

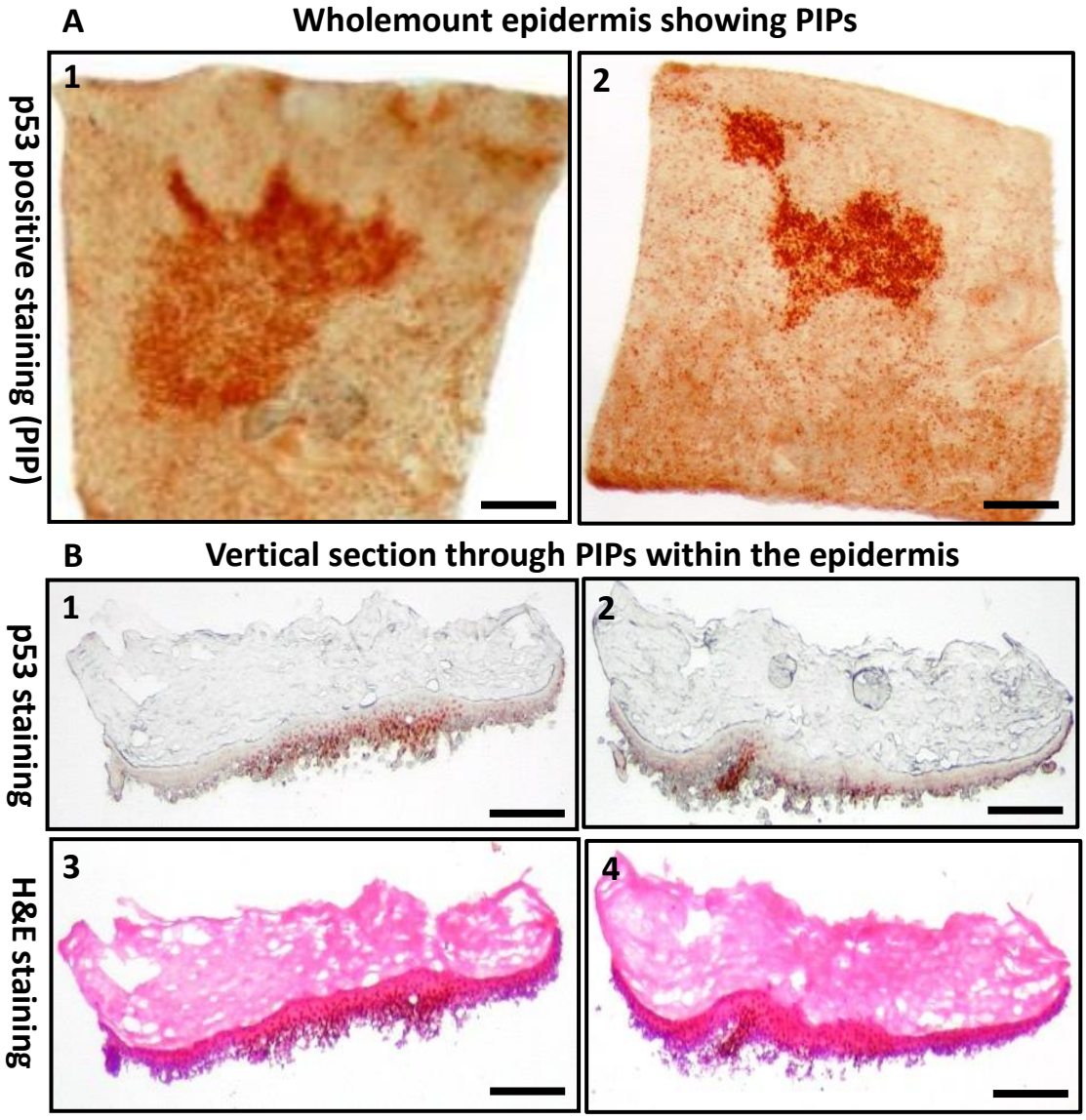


Figure 5.2 The vertical extent of p53 staining within PIPs in epidermal sheets. **A:** Two 4x images (1 – 2) showing PIPs within wholemount epidermal sheets. **B:** Four 4x images of vertical sections through PIPs; the upper row (1 – 2) shows p53 staining within the nucleated part of the epidermal sheets and the second row (3 – 4) shows, on H&E staining that the p53 stains most of the nucleated part of the epidermis with the anuclear stratum corneum above this . Scale bar; 200 µm.

p53 positive nuclei were present in the basal layer and extended into the suprabasal layer in all cases. Although the vertical extent of p53 staining within the PIPs was variable, the extent of the staining within most PIPs was up to the upper part of the stratum spinosum (figure 5.3). (note that stratum granulosum and stratum corneum do not contain nuclei).

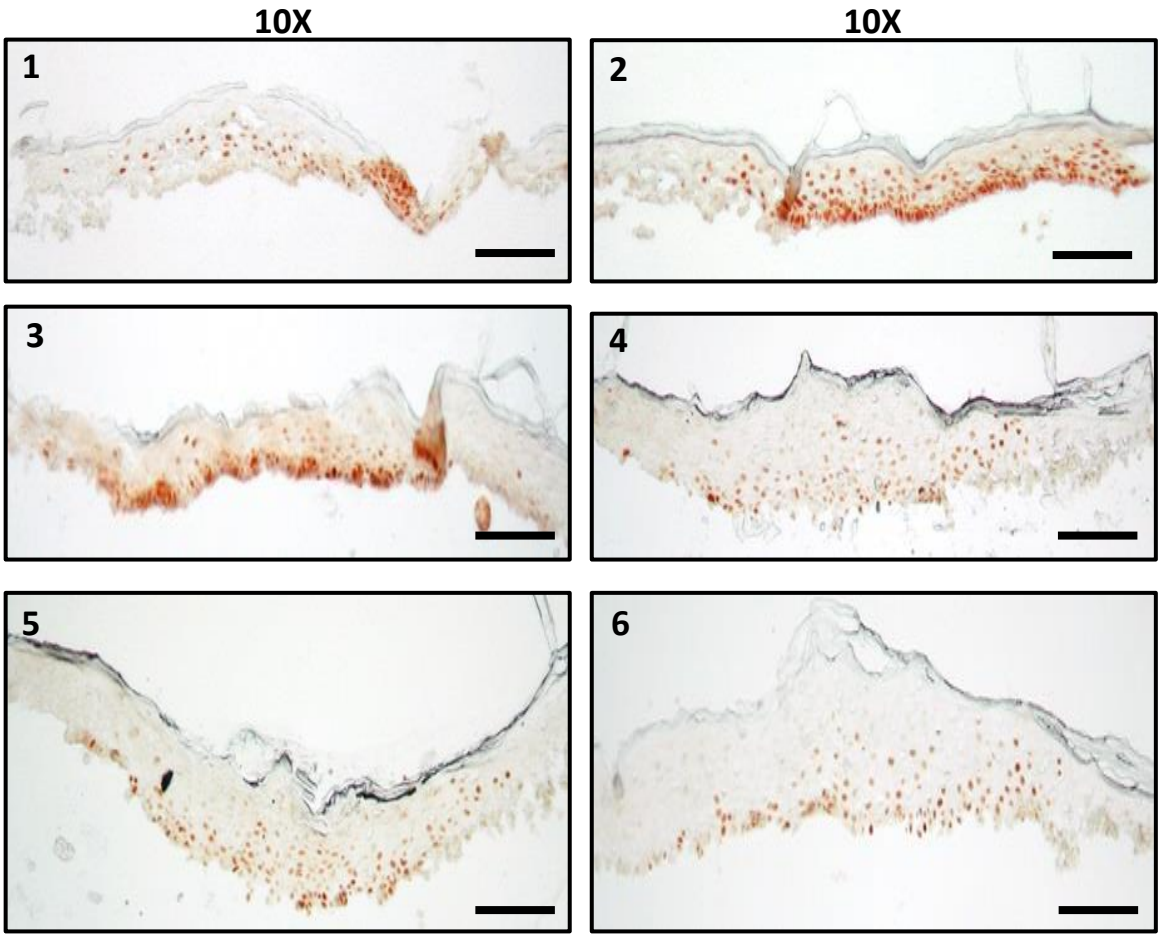


Figure 5.3 Vertical sections through PIPs showing the extent of p53 staining within the epidermal layers. 10x images for vertical sections through the epidermal sheet from 6 different PIPs (1 – 6). In most of cases, the p53 staining extends through all nucleated layers of the epidermis (up to the upper part of stratum spinosum). Scale bar; 50 μm.

After all the previous investigations, PIPs suitable for target enriched sequencing were selected. The details of the 15 subjects which remained for target enriched sequencing data included 8 (53.3 %) males and 7 (46.7 %) females, with subject ages ranging from 67 – 94 years (median 81.5 years). Most of these PIPs were on sun exposed skin areas and were distributed mainly on the face and scalp followed by the upper extremities. The result is illustrated in table 5.3 and appendix table 8.17.

Table 5.3 Details of gender and age of subjects and site of PIPs retained for target enriched sequencing.

	Subject's information		**PIPs**
	Gender		**Number of PIPs**
	Male	**Female**	
Total number	**8**	**7**	**23**
Site of the lesion			
Left side of scalp	1		3
Left temple		1	1
Right temple	1		1
Left cheek		2	1 from one subject and 3 from the other subject
Left pre-auricular	1		1
Left helix	1		4
Right shoulder	1		1
Upper chest		1	1
Left forearm		1	1
Right forearm		1	1
Right hand	1		1
Right hand		1	1
Abdomen	1		1
Shin of the leg	1	1	one per each subject
Age at biopsy	Range (68-94 years), Median 82 years	Range (67-86 years), Median 80 years	

As illustrated in table 5.2, 23 PIPs were selected for target enriched sequencing. DNA extraction was performed on both PIPs and the surrounding non-PIP epidermis (reference DNA). Total DNA yield ranged between 102 – 559.5 ng with median of 244.5 ng in PIPs and between 213 - 870 ng with median of 550.5 ng in the surrounding non-PIP epidermis (figure 5.4 and appendix table 8.18). A Nanodrop spectrophotometer was used to assess DNA purity and the results of the 230/280 and 260/280 ratios are summarised in figure 5.5.

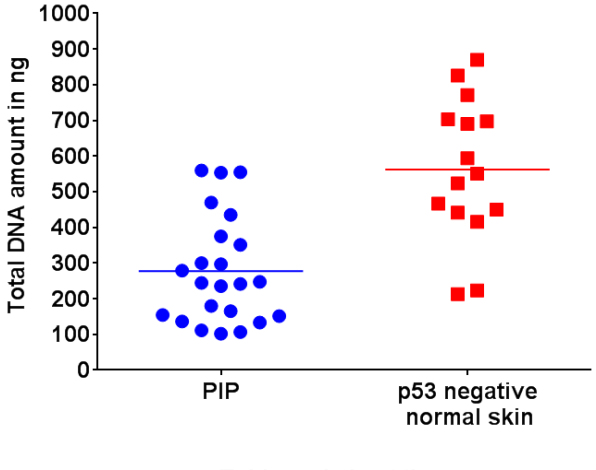


Figure 5.4 Total DNA amount isolated from PIP samples selected for target enriched sequencing. PIPs were dissected under Stereo zoom microscope using straight dissecting needle (Thomas Scientific). DNA was extracted from the dissected PIP and, separately, from the p53 negative epidermis. DNA amount was measured using Qubit 2.0 fluorometer and expressed in ng.

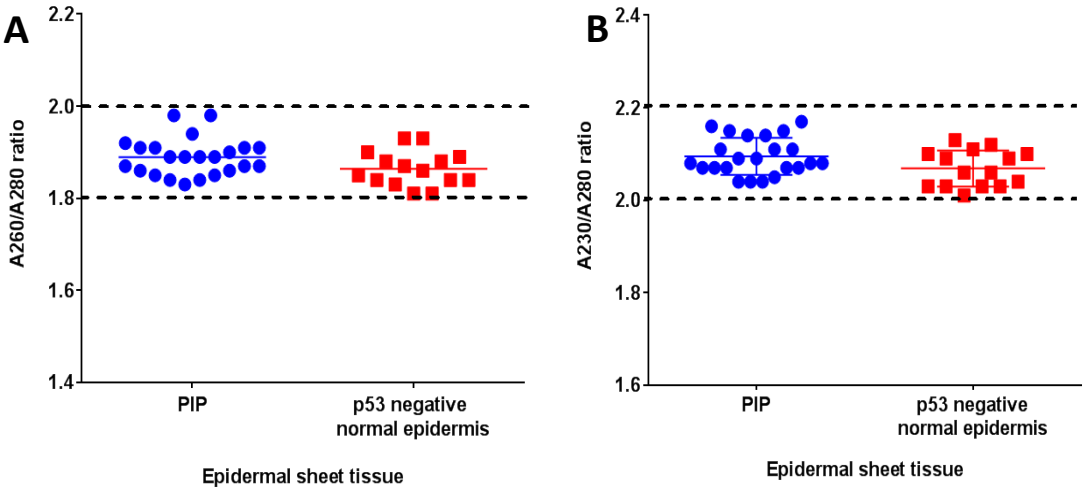


Figure 5.5 DNA quality assessments for the PIPs and corresponding non-PIP epidermis (p53 negative normal epidermis selected for target enriched sequencing. The quality was assessed using Nanodrop spectrophotometer/ UVR absorbance spectra of DNA samples extracted from these tissues. **A**: DNA A260/A280 ratio; values between 1.8 and 2.0 indicate high purity DNA with minimal protein contamination. **B**: DNA A260/A230 ratio values between 2.0 and 2.2 indicate high purity DNA with minimal solvent contamination.

DNA fragmentation was assessed using panel (mutiplex) PCR. All samples (N = 23 PIPs (including multiple PIPs came from some individuals) and 15 non-PIP epidermis) were successful in giving positive PCR results with amplification of DNA fragments up to 600 bp in length. An example of the multiplex PCR results (N = 5) is shown in figure 5.6.

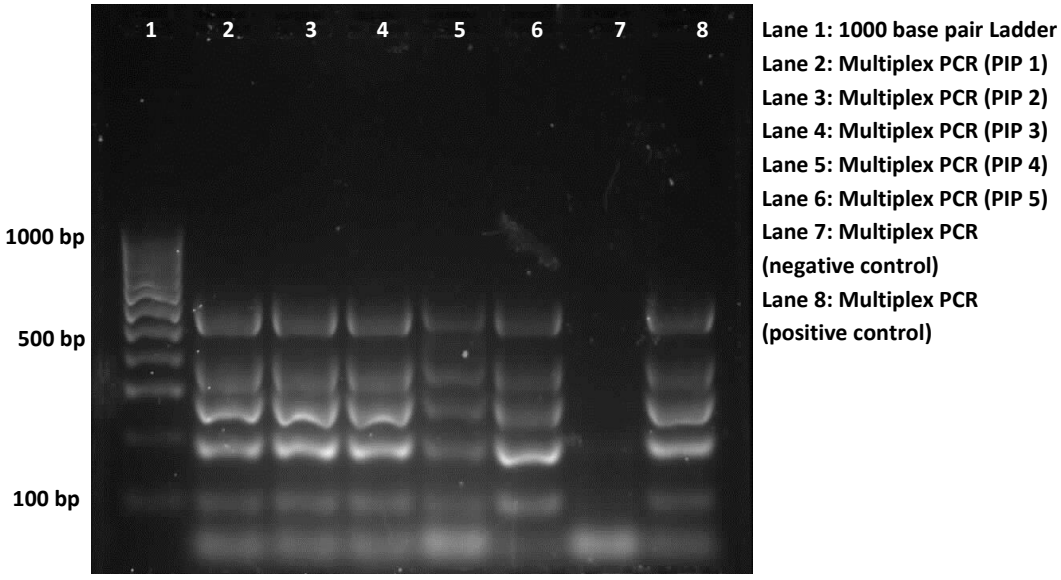


Figure 5.6 Result of agarose gel electrophoresis for panel (multiplex) PCR products. PCR was performed on DNA samples extracted from PIP samples selected for target enriched sequencing using a variety of primers that amplifies DNA to give PCR products of different lengths (100, 200, 300, 400 and 600 bp). Lane 1; DNA ladder with DNA fragments at incremental lengths up to 1,000 bp. Lanes 2-6; DNA samples from PIPs (N = 5). Lane 7; Negative control (no DNA in PCR). Lane 8: Positive control (blood DNA). bp = base pair.

5.4.1 Target enriched sequencing data

5.4.1.1 Primary data analysis

The number of PIPs that fulfilled the quality and quantity criteria and underwent target enriched sequencing was 21 out of the 23 selected PIPs (2 samples were sent for sequencing but were not sequenced as they contained marginally low amounts of DNA when reassessed by the Wellcome Trust Centre for Human Genetics, Oxford using Qubit fluorometer) along with matched p53 negative normal looking epidermis. The analysis was conducted exactly as per that described for the AKs, BDs and cSCCs in chapter 4 and table 5.4 provides data on the PIPs in relation to the data from the AKs, BDs and cSCCs for comparison (i.e. encompasses data from table 4.4 in chapter 4 along with the PIPs results). The highest mean depth of coverage was seen in the PIPs and was equal to 268.8 (range 147.9 – 440.3). Only somatic non-silent (missense, nonsense, frameshift insertion/deletion and splicing) alterations were included within the analysis, and the lowest numbers of these mutations were observed in the PIPs, with the total number of mutations in the PIPs equal to 135 mutations (median of 6 mutations and range 1 - 16 mutations per PIP).

As mentioned previously (section 4.4.1.1), the size of the sequenced genome in the target enriched sequencing was 196 kilobase, thus the mutation burden for the PIPs was 0.03 mutations/kilobase (equal to 30 mutations/megabase) of DNA. As stated in chapter 4, this high mutation burden per megabase of DNA may be due to the fact that the targeted sequencing focussed on genes which were likely to be mutated in these lesions.

Table 5.4 Updated primary results from the target enriched sequencing data. Based on the results that were presented in table 4.4.

	Solitary AK	Solitary BD	AK adjacent to cSCC	BD adjacent to cSCC	cSCC adjacent to AK	cSCC adjacent to BD	PIP
Number of samples	25	29	7	8	7	8	21
Mean fold coverage	179.5 (range 79.7 – 305.4)	166 (range 0.55 – 369.7)	191.4 (range 40.8 – 302.9)	183.6 (range 37.1 – 274.5)	205 (range 72.5 – 313.2)	188.5 (range 51.1 – 269.4)	268.8 (range 147.9– 440.3)
Total number of non-silent mutations	1345	759	256	530	260	517	135
Median number of mutations /sample	25.5	10	14	24	13	16	6
Mean number of mutations /sample	56.04	27.1	36.5	66.5	37.1	64.6	6.9
Range of mutations /sample	6- 296	1-131	2-149	9-216	12-124	3-225	1 - 16
Median mutation burden/ kilobase of DNA	0.13	0.05	0.07	0.12	0.07	0.08	0.03
Median mutation burden/ megabase of DNA	130	50	70	120	70	80	30

By comparing the PIP results with the results of the AKs, BDs and cSCCs (from figure 4.5 in chapter 4), it can be seen that in general the PIPs showed the lowest mutation frequency (silent and non-silent mutations) within the targeted genes (figure 5.7). Although PIPs contained mutations in genes such as *GPR98, NOTCH1*, *HMCN1* and *TP53,* no somatic mutations were detected in the *HRAS* and *NRAS* genes.

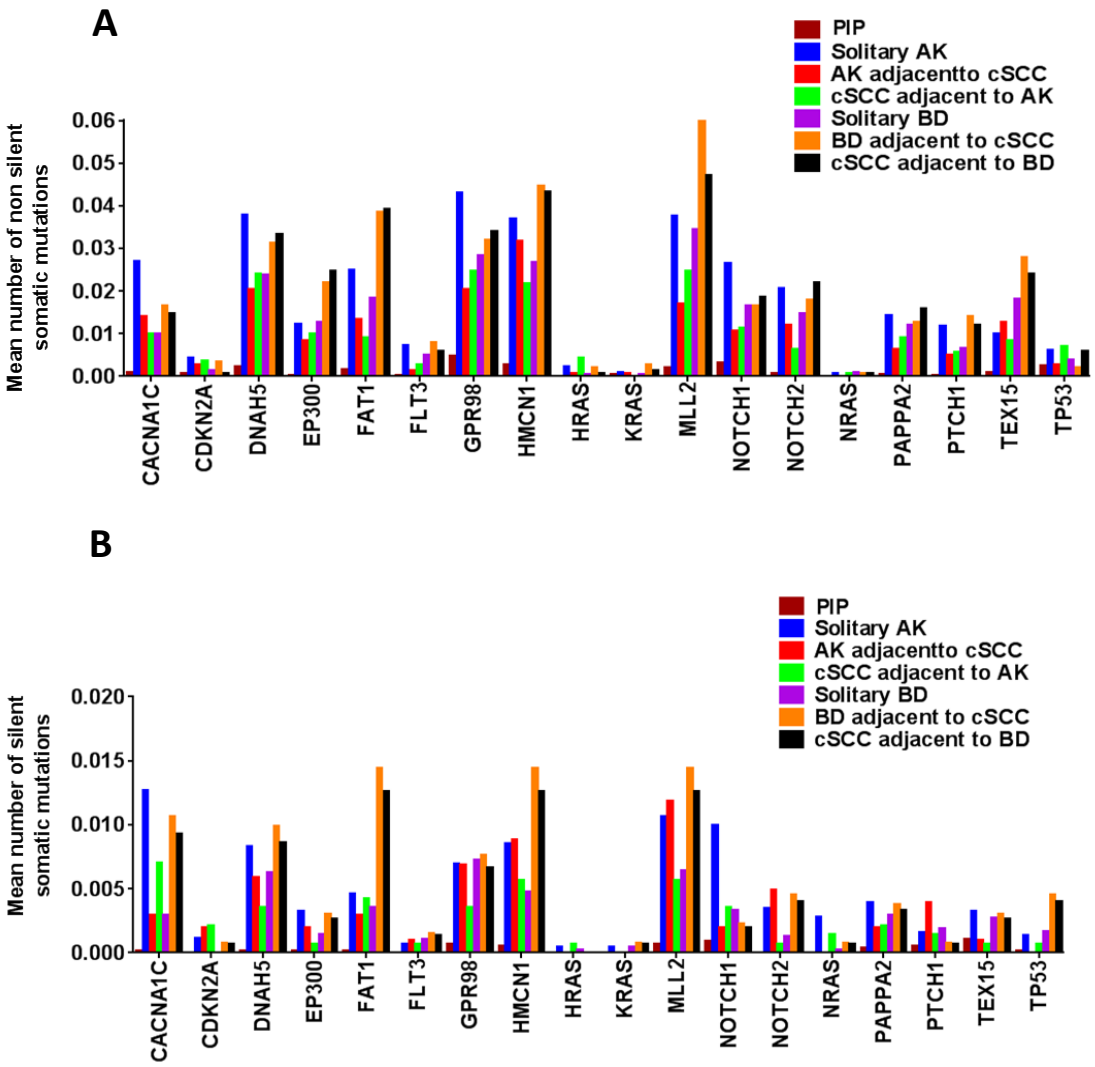


Figure 5.7 Overview of the target enriched sequencing data for PIPs in comparison with the data from AKs, BDs and cSCCs from chapter 4**. A:** Mean number of non-silent somatic mutations. **B:** Mean number of silent somatic mutations within the 18 targeted genes.

5.4.1.2 PIPs variants remaining after the second filtration step

The PIP data was filtered as per that conducted for the AKs, BDs and cSCCs in chapter 4 (sections 4.3.7.3 and 4.4.1.2). As stated in chapter 4, the filteration criteria were selected based on that used in the published studies using NGS on cSCC (Li et al., 2015, Pickering et al., 2014, South et al., 2014). The filtration conditions comprised a minimum mean fold coverage of ≥ 100x, with variants identified as mutations requiring to be present in ≥ 4 reads from the lesion and < 4 reads from the normal skin, and a p value for the somatic changes of ≤ 0.05. Although this filtration resulted in a fewer numbers of PIPs in the final

data set, it generated more robust data which could be interpreted in a way which produced more robust conclusions.

15 of the 21 PIPs remained after the second filteration step and the number of non-silent somatic mutations was reduced, with 69 mutations of the original 135 mutations included in the downstream analysis (median of 4 mutations per PIP, range 1 - 11). As expected from the fact that UVR exposure is the main risk factor for development of PIPs, C > T transition base substitutions (median frequency of 75%, range 33.3 - 100%) were more common than other types of base pair changes. The median frequency of CC > TT mutation was 25%, but were present only in 5 samples of the 15 PIPs. The ratio of non-silent to silent (synonymous) somatic mutation was 3:1. Interestingly, 13 of the 18 targeted genes were mutated within PIPs, with at least 3 genes mutated in 60% (9/15) of PIPs and at least 5 genes mutated in 27% (4/15) of PIPs. *TP53* mutations were noted in 60% (9/15) of PIPs and mutations in other genes such as *GPR98, NOTCH1* and *MLL2* were also seen in 6 or more of the PIPs. Most of the mutations, as seen on the heat map in figure 5.8, were missense mutations. Furthermore, all of the *CDKN2A* and *TP53* mutations identified in the PIPs had been previously reported as alterations in cancers in the Catalogue of Somatic Mutations in Cancer (COSMIC) database. In contrast, other genes showed some (such as *GPRT98* and *NOTCH1)* or no (such as *PTCH1* and *TEX15*) previously reported mutations in COSMIC (figure 5.8).

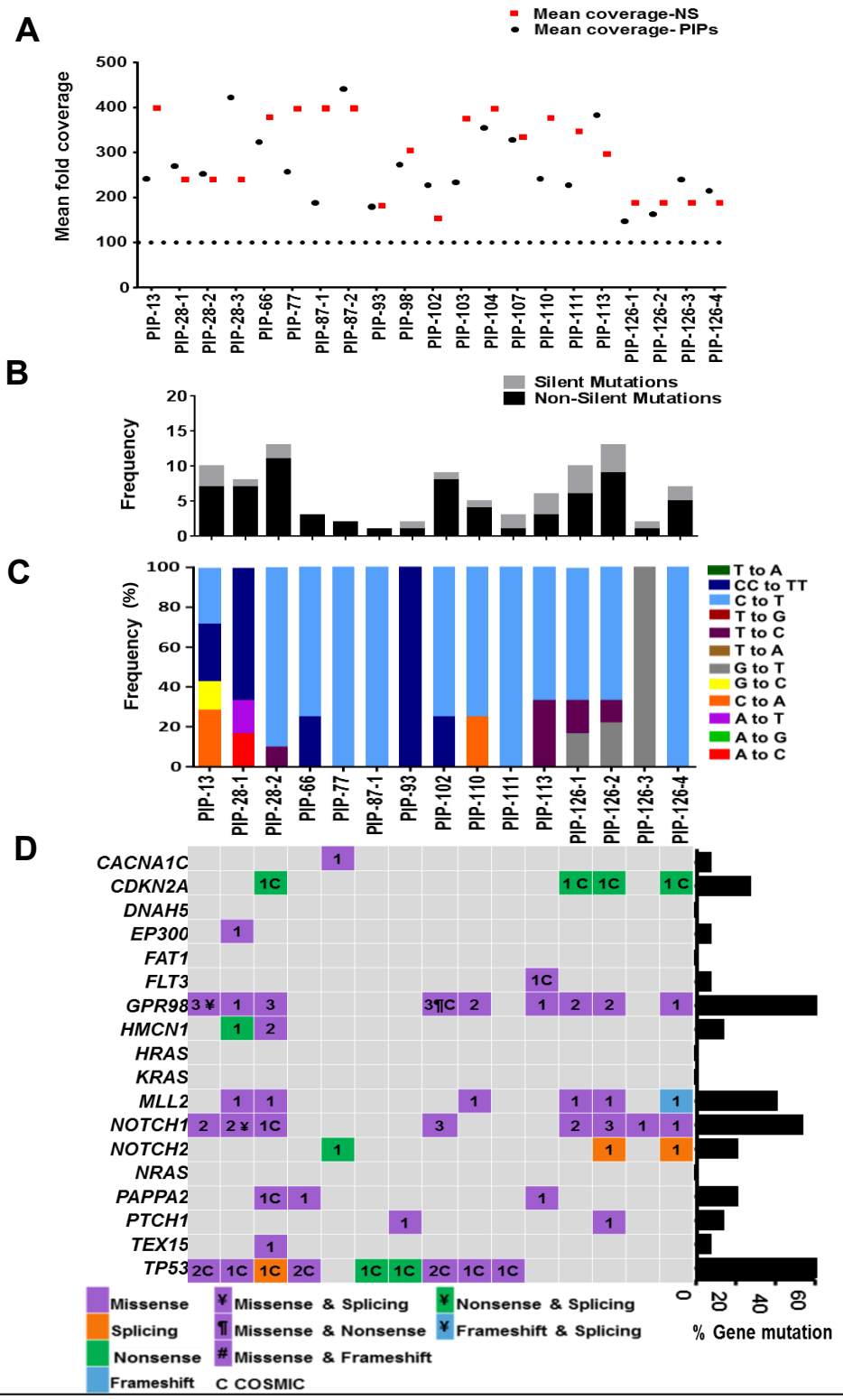


Figure 5.8 Integrated overview on the target enriched sequencing data from PIPs ; part A shows results from the 21 PIPs which underwent target enriched sequencing, whereas parts B – D show results from the 15 PIPs remaining after the second round of filtration. **A:** The mean fold coverage across the 21 PIPs. **B:** The frequency of silent and non-silent genetic alterations within the PIPs is plotted on the top of the heatmap. **C:** The frequency of different base pair changes within the PIPs showing the highest frequency of mutation to be C > T changes. **D:** Heatmap representation of the different types of mutations within the target genes within different PIPs. The corresponding bar graph on the right of each heatmap shows the mutation frequency represented as the percentage of gene mutation within the samples. The numbers within the heatmap indicate the number of mutation/gene/sample with the different colours each indicating a different type of mutation. Note that in B – D, each of the columns represent the same PIPs.

Using (SIFT) (www.sift.bii.a-star.edu.sg), the number of non-silent mutations with predicted deleterious effect within each gene was calculated and the results summarised in table 5.5. Many of the genes showed high number of non-silent mutations with deleterious effect (*TP53, NOTCH1* and *GPR98* harboured the highest number of deleterious mutations). Overall, 33 of the 69 detected mutations within PIPs (48%) were predicted to be deleterious.

Table 5.5 Non-silent mutations with predicted deleterious effect within PIPs. The effect of the mutations was predicated according to the result of SIFT and the number of the deleterious mutations is plotted per the total number of non-silent mutation per gene

	CACNA1C	*CDKN2A*	*EP300*	*FLT3*	*GPR98*	*HMCN1*	*MLL2*	*NOTCH1*	*NOTCH2*	*PAPPA2*	*PTCH1*	*TEX15*	*TP53*
The proportion of mutations with deleterious effect	**0/1**	**1/4**	**1/1**	**1/1**	**10/18**	**0/3**	**4/6**	**5/15**	**1/4**	**2/3**	**0/2**	**1/1**	**8/12**

5.4.1.3 Clonal evolution of mutations within PIPs

Clonality analysis was conducted on the PIPs in the same way as it was conducted on the AKs, BDs and cSCCs in chapter 4 (section 4.4.1.2.4). Only 4 genes (*CDKN2A*, *CACNA1C*, *MLL2* and *GPR98*) harboured clonal mutations within a proportion of the PIPs, with the majority of the mutations were sub-clonal in the overall group of PIPs (figure 5.9). In contrast to AKs, BDs and cSCCs, all of the detected *TP53* mutations within the PIPs (N = 12) were subclonal, indicating that *TP53* mutation was unlikely to have been the initiating event in each of these PIPs.

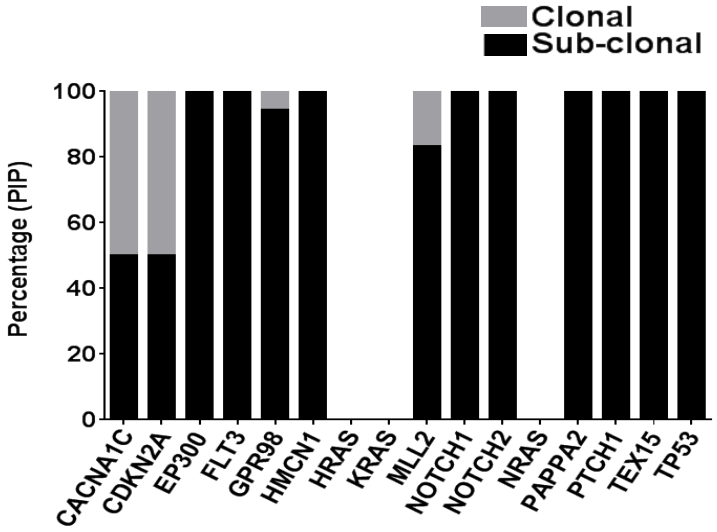


Figure 5.9 Clonality assessments of the somatic variants within the mutated genes in the PIPs. Genes are displayed from left to right in alphabetical order. The classification of variants as clonal or subclonal was performed using a binomial distribution based on the alternative allele read count, the total read count from lesional cells and an expected variant allelic fraction (VAF) of 0.45. The result is expressed as percentage. PIPs = p53 immunopositive patches.

5.4.1.4 PIPs share several mutations with precancerous lesions (AK and BD) and cSCCs

To determine whether there was an obvious difference in the genes which were mutated in the PIPs and those which were mutated in the later precancerous lesions (AKs and BDs, and/or cSCCs), the results of the target enriched sequencing of the PIPs were compared with those of the AKs, BDs and cSCCs (as described in chapter 4, table 4.7). The comparison showed that all of the genes mutated in the PIPs were also mutated in the AKs and BDs, as well as in the cSCCs (table 5.6).

Table 5.6 Comparison of the percentage of number of samples with mutations in each of the 18 targeted genes within all lesional groups. The groups were as follows, all AKs combined (i.e. solitary AKs and AKs adjacent to cSCC), all BDs combined (i.e. solitary BDs and BDs adjacent to cSCC), cSCCs (adjacent to AKs or BDs) and PIPs .

	CACNA1C	*CDKN2A*	*EP300*	*FLT3*	*GPR98*	*HMCN1*	*HRAS*	*KRAS*	*MLL2*	*NOTCH1*	*NOTCH2*	*NRAS*	*PAPPA2*	*PTCH1*	*TEX15*	*TP53*	*FAT1*	*DNAH5*
PIP	7	27	7	7	60	13	0	0	40	53	20	0	20	7	7	60	0	0
AK	71	33	21	29	58	63	13	4	79	71	50	0	54	29	46	42	0	0
BD	39	22	30	13	39	48	4	4	48	65	43	9	35	17	26	39	0	0
cSCC	33	25	17	8	58	42	25	8	75	58	33	8	58	25	50	25	0	0

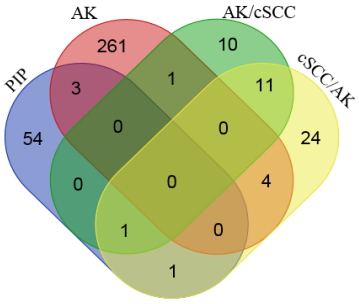
Some of the detected mutations were identical between the different lesions (i.e. the same amino acid alteration in the same gene). The highest number of shared mutations is seen between the adjacent lesions, with 12 mutations shared between adjacent AKs and

cSCCs and 10 mutations shared between adjacent BDs and cSCCs (figure 5.10). Solitary lesions (AKs and BDs) also shared some mutations, either between each other or with cSCCs (figure 5.10), for example, one solitary AK and a BD adjacent to cSCCs as well as the corresponding cSCC adjacent to the BD contained a G13D mutation in *HRAS*. Moreover, the PIPs shared mutations (some of which are expected to be deleterious (nonsense)) with the AKs and/or BDs and/or cSCCs. From the data in figure 5.10, it can be seen that 8 mutations (affecting *CDKN2A, NOTCH1, NOTCH2, PAPPA2* and *TP53*) were common between PIPs and AKs, BDs or cSCCs, for example, an identical W110* nonsense mutation in *CDKN2A* was observed in one PIP and one case of BD, whereas a Q50* nonsense mutation in the same gene was noted in a PIP, AK and cSCC. In addition, an identical *NOTCH2* nonsense mutation (Q1560*) was detected in a PIP and BD, while a non-silent change at codon 57 in the *PAPPA2* gene was seen which resulted in a missense mutation in a PIP and a frameshift mutation in a BD (figure 5.10).

A

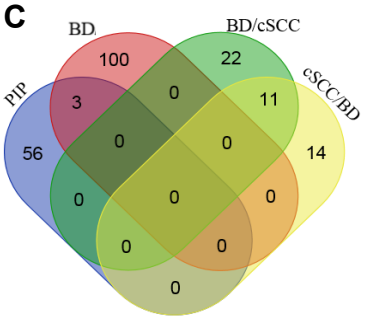
	PIP	AK	Adjacent lesions		BD	Adjacent lesions	
			AK	cSCC		BD	cSCC
CACNA1C			C1846W	C1846W			
CDKN2A	Q50*		Q50*	Q50*			
	W110*				W110*		
GPR98		H277Y		H277Y			
			P1481L	P1481L		E3733K	E3733K
HRAS		G13D				G13D	G13D
KRAS		G12D		G12D			
MLL2		V659E	V659E			R5448*	R5448*
		G4596fs					G4596fs
		G4596*	E767G		E767G	G4596*	
			Q425*	Q425*			
		V658E		V658E		E3569K	E3569K
		P2198L			P2198L		
NOTCH1	H599Y	H599Y	P1770S	P1770S		P1381fs	P1381fs
	G384S	G384S				Q1810*	Q1810*
NOTCH2	Q1560*			Q1560*		E798*	E798*
PAPPA2	R57Q		E296K	E296K	R57*		
	E1443K	E1443K					
PTCH1			P1135L	P1135L		Y870*	Y870*
			E175K	E175K		E511K	E511K
			K237K	K237K			
						S2484F	S2484F
TP53	R116W	H47Y		H47Y	R116W		
			P359L	P359L			
		R210*	S127F	S127F	R210*		
		R210P	P359S	P359S			

B



NOTCH1: H599Y, G384S
PAPPA2: E1443K
NOTCH2: Q1560*
CDKN2A: Q50*

C



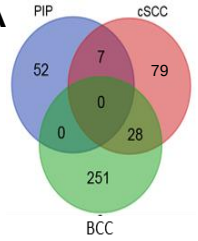
TP53: R116W
CDKN2A: W110*
PAPPA2: R57Q

Figure 5.10 Shared mutations between different lesions. **A:** The position of each shared mutation and the altered amino acid on the resultant protein of all samples which were investigated by target enriched sequencing. **B:** Venn diagram highlights the shared mutations between PIP, AK, AK adjacent to cSCC and cSCC adjacent to AK. **C:** Venn diagram highlights the shared mutations between PIP, BD, BD adjacent to cSCC and cSCC adjacent to BD. PIP = p53 immunopositive patch. cSCC = cutaneous squamous cell carcinoma. AK = actinic keratosis. BD = Bowen's disease.

Comparison of PIPs in this study with cSCCs in the published literature/databases (Chitsazzadeh et al., 2016, Durinck et al., 2011, Lee et al., 2014, Li et al., 2015, Pickering et al., 2014, South et al., 2014) also identified 7 alterations in 4 genes (*PAPPA2, NOTCH2, CDKN2A* and *TP53)* which were identical in both lesions. By contrast, no identical mutations were seen in the PIPs in this study and basal cell carcinomas (including those investigated by WES) in the published literature/databases (Boldrini et al., 2003, Bolshakov et al., 2003, Bonilla et al., 2016, Evans et al., 2000, Harwood et al., 2008, Heitzer et al., 2007, Huang et al., 2013, Jayaraman et al., 2014, Kim et al., 2002, Pacifico et al., 2008, Rady et al., 1992, Reifenberger et al., 2005, Reifenberger et al., 1998, Soufir et al., 1999, Van Der Riet et al., 1994, Van Der Schroeff et al., 1990, Weihrauch et al., 2002, Wilke et al., 1993, Zaravinos et al., 2010) (figure 5.11).

Due to the limited number of genetic studies on AKs and BDs previously, it was only possible to compare mutations in *TP53* and *CDKN2A* in PIPs and these precancerous lesions. However, the results showed 3 identical *TP53* mutations between PIPs in this study and AKs/BDs in the published literature (Campbell et al., 1993b, Chitsazzadeh et al., 2016, Nelson et al., 1994, Nindl et al., 2007, Pacifico et al., 2008, Park et al., 1996, Soufir et al., 1999, Taguchi et al., 1998, Takata et al., 1997, Zaravinos et al., 2010, Ziegler et al., 1994) (figure 5.11). Comparison of mutations in PIPs and “normal” skin which had been investigated in this thesis by WES as controls for the AKs in chapter 3 (and which are likely to have been chronically sun-exposed) and from normal chronically sun-exposed skin in the literature/databases (Chitsazzadeh et al., 2016, Martincorena et al., 2015) indicated that there were 5 identical mutations in *NOTCH1* and *NOTCH2* between the PIPs and normal chronically sun-exposed skin (figure 5.11).

A

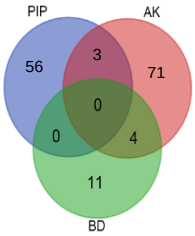


PAPPA2: R664K[1], E1443KL[2]
NOTCH2: P460S[3], P460L[1,3]
CDKN2A: W110* [4], Q50* [1,3,5]

B

BCC (%) gene mutation	Boldrini et al., 2003¥	Bolshakov et al., 2003¥	Bonilla et al., 2016¶,‡	Evans et al., 2000¥	Harwood et al., 2008¥	Heitzer et al., 2007¥	Huang et al., 2012¥	Jayaraman et al., 2014¶	Kim et al., 2002¥	Pacifico et al., 2008¥	Rady et al., 1992¥	Reifenberger et al., 1998¥	Reifenberger et al., 2005¥	Soufir et al., 1999¥	Van Der Riet et al., 1994¥	Welhrauch et al., 2002¥
PTCH1			73	20.4	55	48	54.8	75	40			25.8	67			
TP53	15.7	47.3	61					66	33	41.1	50		40		44	58
CDKN2A			5							35.2				3.5		
NOTCH1			40					50								
NOTCH2			26					66								

C



TP53: P91L [6], P91S [6], R248W [8]

D

AK (%) gene mutation	Nelson et al., 1994¥	Nindl et al., 2007¥	Pacifico et al., 2008¥	Park et al., 1996 ¥	Soufir et al., 1999¥	Taguchi et al., 1998¥	Ziegler et al., 1994¥	Chitsazzadeh et al., 2016¶
TP53	53	6.7		28	100	30	60	100
CDKN2A		2.7	50		100			
FAT1								71
MLL2								71
NOTCH1								71
NOTCH2								57

E

BD (%) gene mutation	Campbell et al., 1993¥	Takata et al., 1997¥
TP53	40	31

F

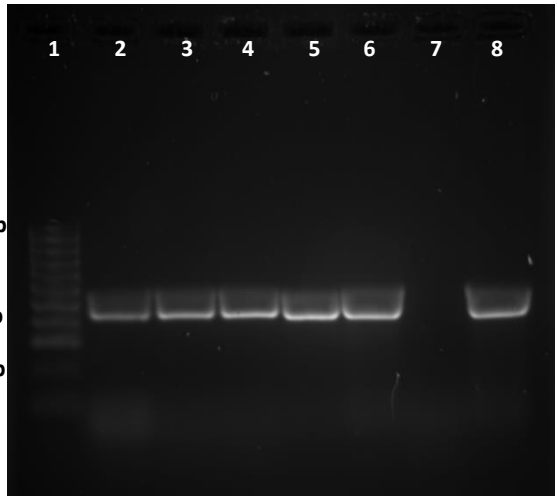
NOTCH1: N325K [7], R365C [7]
NOTCH2: P460S [7], P460L [7], Q1560* [7]

Figure 5.11 Shared mutations between PIPs in this study and published results on cSCCs, BCCs, AKs, BDs and normal chronically sun-exposed skin. **A&C:** Venn diagram comparing the somatic mutations within the 13 mutated genes in PIPs and reported mutations within **(A)** cSCCs and BCCs and **(C)** AKs and BDs from the literature. **B,D &E:** Percentage of mutations in genes which were mutated in PIPs and in **(B)** BCCs, **(D)** AKs and **(E)** BD. **F:** Venn diagram highlighting the shared mutations between PIPs and normal skin of the WES AK data (N = 5) [NS(1)] and normal skin from Martincorena et al. (2015) and Chitsazzadeh et al. (2016) [NS(2)]. References; 1- Pickering et al., 2014, 2- Durinck et al., 2011, 3- Li et al., 2015, 4- South et al., 2014, 5- Ping et al., 2001, 6- Ziegler et al., 1994, 7- Martincorena et al., 2015, 8- Chitsazzadeh et al., 2016. The ¥ sign indicates the experiments performed using Sanger sequencing, ¶ indicates whole exome sequencing based research and ‡ indicates target enriched sequencing based results. PIP = p53 immunopositive patch. cSCC = cutaneous squamous cell carcinoma. BCC = basal cell carcinoma. AK = actinic keratosis. BD = Bowen's disease.

5.4.2 *KNSTRN* Sanger sequencing analysis

Sanger sequencing for the first exon of *KNSTRN* was performed on DNA extracted from the 21 sequenced PIPs (with the p53 negative adjacent skin for each PIP sample) selected for the target enriched sequencing. As was the case for the *KNSTRN* sequencing in chapter 4, SeqBuilder (v10.0.1.3) and Seq Scanner (v2.0) software (Applied Biosystems) were used to analyse the results, and only those bases with a quality score (Phred score) ≥ 20 were included in this analysis. *KNSTRN* mutations were detected in 6 of the 21 PIPs on both the sense and antisense strands of DNA. Similar to the other lesions in this thesis in which *KNSTRN* was sequenced (i.e. AKs, BDs and cSCCs in chapter4), the mutation was the p.Ala40Glu alteration, but in the PIPs all (N = 6) were heterozygous alterations (figure 5.12).

A



Lane 1: 1000 base pair Ladder
Lane 2: *KNSTRN* PCR product (PIP-1)
Lane 3: *KNSTRN* PCR product (PIP-2)
Lane 4: *KNSTRN* PCR product (PIP-3)
Lane 5: *KNSTRN* PCR product (PIP-4)
Lane 6: *KNSTRN* PCR product (PIP-5)
Lane 7: *KNSTRN* PCR product (negative control)
Lane 8: *KNSTRN* PCR product (positive control)

B Information on *KNSTRN* gene mutation detected in PIP samples.

Number of PIPs	Detected variant on lesional DNA	Base pair change	Variant position on DNA	Protein change	cDNA change	SIFT	PolyPhen
6	GCG (C>A)	Missense	15:40382954-40382954	p.Ala40Glu	c.119C>A	Deleterious	Probably damaging

C

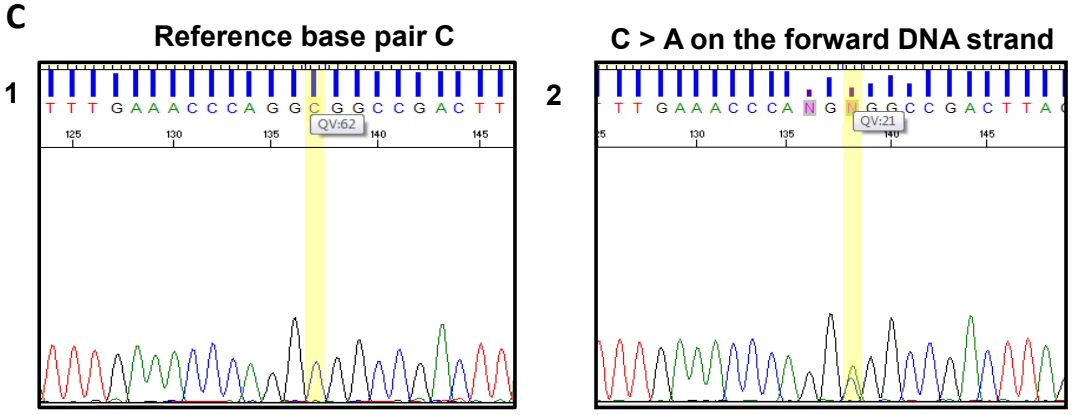


Figure 5.12 *KNSTRN* mutations in PIPs. **A:** Agarose gel electrophoresis results from PCR of DNA samples extracted from 5 PIPs using a primer pair that amplified *KNSTRN* exon 1 to give a 317 bp product. Lane1: DNA ladder with DNA fragments of different lengths up to 1,000 bp. Lane 2-6; PCR product from 5 PIPs (labelled as 1, 2, 3, 4 and 5). Lane7: Negative control (no DNA in PCR). Lane 8: Positive control (blood DNA). **B:** Information on the detected mutation using the Variant Effect Predictor tool on the Ensembl webpage (Ensembl) (www.ensembl.org/info/docs/tools/vep/). **C:** Sequencing analysis using SeqScanner (v2.0) software showing **(1)** the reference DNA with a C base (highlighted within yellow band) and **(2)** a C > A change on the forward DNA strand from a PIP (also highlighted by yellow band). bp = base pair.

5.5 Discussion

Prior to this study, there has been limited study of the genetic alterations within PIPs. Previous research has generally concentrated on *TP53* mutational status within PIPs (Jonason et al., 1996, Kramata et al., 2005, Robinson et al., 2010), but one study has investigated loss of heterozygosity in these lesions (finding allelic loss predominantly on 9q in 28.3% of cases) (Tabata et al., 1999). Based on the hypothesis that PIPs may represent early clinically invisible precancerous lesions which may develop into cSCC, the aim of this study was to investigate for genetic changes in PIPs using a target enriched sequencing approach for genes which were identified as mutated in AK, BD and cSCC lesions.

As per the case in many of the previous publication on PIPs, *TP53* was frequently mutated, but was not mutated in all PIPs in this study. Moreover, the results showed that *TP53* mutation is not the only mutational event, and that other genes are also mutated within these patches. The most frequently mutated genes in the PIPs were *TP53* and *GPR98* (each 60%). *GPR98* (i.e. G protein-coupled receptor 98) is one of the most frequently mutated G protein-coupled receptors in cancer (O'Hayre et al., 2013), with genes within the G protein superfamily classified as oncogenes and reported to be mutated in around 20% of all human cancers (Kan et al., 2010). The *GPR98* mutations in the PIPs were mainly missense, resulting in amino acid alterations. Indeed, according to SIFT software, 48% of the mutations in all 13 genes which were mutated in the PIPs were considered to alter function. *TP53, NOTCH1* and *GPR98* showed the highest number of deleterious mutations that were located mainly in the p53 DNA binding domain, Notch1 EGF like repeat domain and Gpr98 Calx-beta domain of the resultant proteins.

NOTCH1 was also within the top mutated genes in PIPs, with the frequency of *NOTCH1* gene mutations within PIPs (53%) similar to that previously reported in cSCCs (Li et al., 2015, Pickering et al., 2014, South et al., 2014). A recent study in 2015 by Martincorena and his colleagues identified *NOTCH1* as the most frequently mutated gene in chronically sun-exposed eyelid skin, In that study, they estimated that 14 to 21% of normal looking skin epidermal cells carried *NOTCH1* mutations and that the frequency of *NOTCH1* mutations within 5 cm^2 of normal looking skin is higher than the frequency of *NOTCH1* mutations within 5,000 cancers reported by TCGA (The Cancer Genome Atlas)

(Martincorena et al., 2015). What this current study on PIPs suggests is that many of these *NOTCH1* mutations in skin chronically exposed to sunlight are probably present within PIPs, but it is not possible to exclude the existence of "non-PIP" patches with *NOTCH1* mutations (i.e. p53 immunonegative patches of *NOTCH1* mutated keratinocytes) on the basis of the current study. In addition, Chitsazzadeh et al. (2016) showed *NOTCH1* mutations within normal sun exposed skin samples together with *MLL2*, *TP53* and *FAT1* mutations.

Assessment of PIPs clonality revealed clonal mutations within only 4 of the 13 mutated genes. However, none of the detected *TP53* mutations were clonal within these PIPs. This suggests *TP53* mutations are a later event in the process of PIP development / evolution. Indeed, the fact that PIPs are considered clonal lesions, and their identification as a discrete lesion when viewed in wholemount p53 immunostained epidermis, as well as the ability to pick up clonal mutations in a number of the PIPs where *TP53* mutations were subclonal in this thesis, all suggest that the *TP53* mutation was not the initiating event in PIP development. In the PIPs where *CDKN2A*, *CACNA1C*, *MLL2* and *GPR98* were clonal, it is possible that the mutation of these genes was the initiating genetic event. However, it is also theoretically possible that another genetic event that might have occurred in the original precursor cell at the same time as, or just before, that event could have been the initiating stimulus for clonal growth of the PIPs. The PIPs were originally detected by their increased p53 protein content by virtue of their immunopositivity, but it is not surprising that *TP53* was not mutated in all PIPs because elevated levels of p53 protein can occur as a result of protein stabilisation for other reasons, e.g. DNA damage (Chikayama et al., 2000, Speidel, 2015, Younger et al., 2015). It should also be highlighted that this study focussed on sequencing all exons of the *TP53* (as was the case for all 18 genes in the targeted sequencing) and would therefore not have identified mutations in the promoter region of *TP53* that might have accounted for over-expression of the protein.

There is evidence, mainly from studies in mice, to suggest that PIPs can develop into cSCC (Berg et al., 1996, De Graaf et al., 2008, Kramata et al., 2005, Melnikova et al., 2005, Rebel et al., 2005, Van Kranen et al., 2005). It is unclear whether PIPs can develop into AKs and BDs, but the observation in this study that similar genes were mutated in PIPs, AKs, BDs and cSCCs, as well as the finding of similar mutations in certain genes in the PIPs

and these other lesions, suggest that PIPs are precursors of these other later precancerous lesions and of cSCC in humans. Admittedly, that does not mean that each of the PIPs would have ultimately developed into a cSCC if the skin samples had not been taken from the patients, because there are much greater numbers of PIPs than cSCCs in skin (Rebel et al., 2005). It is not yet clear why some PIPs would progress to skin cancer, although it is presumed that further genetic changes would be required for this to happen. Alternatively, changes in skin immunity may be responsible for progression to later precancers and to cSCC, but further research will be required to elucidate this.

The relationship between PIPs and basal cell cancers is not clear. Genetically engineered mice have demonstrated that cells present in the hair follicle and interfollicular epidermis can give rise to BCCs (Kasper et al., 2012), and a human study has documented that 36% of basal cell cancers arose in lesions previously clinically diagnosed as AKs (Criscione et al., 2009). However, 5 of the same genes (*TP53, CDKN2A, NOTCH1, NOCT2* and *PTCH1*) were mutated in the PIPs in this study and in BCCs in the published literature (Boldrini et al., 2003, Bolshakov et al., 2003, Bonilla et al., 2016, Evans et al., 2000, Harwood et al., 2008, Heitzer et al., 2007, Huang et al., 2013, Jayaraman et al., 2014, Kim et al., 2002, Reifenberger et al., 2005, Reifenberger et al., 1998, Soufir et al., 1999, Van Der Riet et al., 1994, Van Der Schroeff et al., 1990, Weihrauch et al., 2002, Wilke et al., 1993, Zaravinos et al., 2010) which taken in the context of the previously reported loss of heterozygosity of the chromosome 9q microsatellite marker D9S197 in PIPs (Tabata et al., 1999) suggests that PIPs may have the potential to develop into BCCs as well as into cSCCs. Again, further work will be necessary to determine whether this is the case.

Chapter 6: General discussion

6.1 General discussion

Non Melanoma Skin Cancer (NMSC) is the most common human cancer worldwide and includes BCC and cSCC (Lomas et al., 2012). While BCC is classified as a local malignant tumour that rarely metastasises (Jacobs et al., 1982, McCusker et al., 2014), cSCC has the ability to metastasise (Brantsch et al., 2008). Although cSCC is associated with low mortality overall (Rose et al., 2013), its high incidence, requirement for surgical treatment, and the need to follow up substantial numbers of patients with cSCC means that the cost of management associated with cSCC is high. For example, according to USA statistics, the treatment cost is estimated to be $500 million per annum and an additional $2 billion represents the economic impact of the disease, including costs associated with lost productivity (Chitsazzadeh et al., 2013).

cSCC can grow as a *de novo* lesion or on top of precancerous skin conditions (Yanofsky et al., 2011). Potential precancerous skin conditions can be visible lesions such as [AKs (Alam and Ratner, 2001, Ko, 2010, Rigel and Stein-Gold, 2013) and BD (Reizner et al., 1994)] and invisible lesions such as PIPs (Jonason et al., 1996, Rebel et al., 2005, Zhang et al., 2001b) that can only be detected by staining the epidermal sheet for p53 protein.

Chronic UVR exposure is the main environmental risk factor associated with the development of cSCC as well as potential precancerous skin conditions (De Gruijl and Rebel, 2008). One of the UVR associated mechanisms in cancer development is the induction of UVR associated photoproducts that results in C > T or CC > TT changes at dipyrimidine sites (Brash, 2015).

Until the recent development of NGS sequencing, only a few candidate genes (*TP53, CDKN2A, RAS* gene family and *PTCH1*) had been assessed for base pair changes within NMSC lesions (Boldrini et al., 2003, Bolshakov et al., 2003, Brash et al., 1996, Caldeira et al., 2004, Evans et al., 2000, Forbes et al., 2011, Harwood et al., 2008, Heitzer et al., 2007, Huang et al., 2013, Kim et al., 2002, Kubo et al., 1994, Nelson et al., 1994, Oberholzer et al., 2012, Pacifico et al., 2008, Pierceall et al., 1992, Rady et al., 1992, Reifenberger et al., 2005, Reifenberger et al., 1998, Soufir et al., 1999, Van Der Riet et al., 1994, Van Der Schroeff et al., 1990, Van Kranen et al., 1995, Weihrauch et al., 2002, Zaravinos et al., 2010, Ziegler et al., 1994). More recently, the entire coding region of nuclear DNA or of multiple genes has been assessed using WES or targeted sequencing techniques

respectively (Bonilla et al., 2016, Chitsazzadeh et al., 2016, Jayaraman et al., 2014, Li et al., 2015, Pickering et al., 2014, South et al., 2014). Both BCC and cSCC have been found to have a high mutation burden, with BCC having the highest reported mutation rate to date (75 mutations /megabase) (Bonilla et al., 2016).

A limited number of genetic studies have been performed on AKs and BDs, with *TP53, CDKN2A* and *RAS* gene family mutations identified using Sanger sequencing of AK and/or BD lesions (Campbell et al., 1993b, Kanellou et al., 2008, Nelson et al., 1994, Nindl et al., 2007, Pacifico et al., 2008, Park et al., 1996, Soufir et al., 1999, Taguchi et al., 1998, Taguchi et al., 1994, Takata et al., 1997, Zaravinos et al., 2010, Ziegler et al., 1994). Before this current study was initiated, there were no published studies reporting WES analysis of AKs (or of BDs), however, a recent study investigating WES of 9 AKs (and reporting on 7 AKs in their final analysis) was published at the end of August 2016 after the WES and targeted sequencing in the current study was completed. Given that when this study commenced, there was no WES data available on AKs, the first aim of this study was to investigate the mutational landscape of AKs using WES to identify those genes mutated in AKs that had also been previously reported mutated in cSCC, and which might therefore comprise genes that would be more likely to be mutated in PIPs.

The WES results showed a massive mutation load within the AK samples (N = 5) with a median of 34.5 mutations/megabase of exonic DNA. This showed that AK ranks 3rd after BCC (Bonilla et al., 2016, Jayaraman et al., 2014) and cSCC (Durinck et al., 2011, South et al., 2014) in the terms of the burden of mutations within these lesions. Interestingly, this high mutation load is greater than that seen in melanoma, despite melanoma being a more aggressive skin lesion (Hodis et al., 2012). In the more recent study on AKs by Chitsazzadeh et al. (2016), they reported an average mutation rate of 18.5 mutations/megabase in AKs. The lower estimate of the mutation burden in AKs observed by Chitsazzadeh et al. (2016) than that observed in this thesis, may have resulted from contamination with normal cells (including normal keratinocytes and immune cells) in the Chitsazzedah samples because they did not microdissect and isolate DNA from the dysplastic keratinocytes within their samples, but seem to have used the entire piece of clinical tissue for extraction of DNA. By contrast, within this thesis, sections were assessed carefully using various staining procedures and the AK was dissected using the LCM

microscope to ensure that DNA was obtained from the dysplastic tissue. Thus the DNA isolated from the AKs in this study was likely to be of higher purity. Furthermore, within this study, despite the challenges of obtaining DNA with sufficient amount and quality for WES, the advantage of the use of numerous FFPE sections during the LCM process helped ensure that the lesions were true AKs rather than early cSCCs with no evidence of invasion of the basement membrane. Regardless of the differences in the AK mutation burden between the two studies (Chitsazzedah et al., 2016 and the current study), both studies identified that the mutation load in AKs is still higher than that reported within other malignant tumours such as pancreatic, lung and ovarian cancers (Kandoth et al., 2013, Vogelstein et al., 2013). As the skin, gastrointestinal and respiratory system are in direct contact with the external environment, the high mutation rate within lesions originating in these organs has been explained by the continuous exposure of these organs to different mutagens (Vogelstein et al., 2013). However, this view does not explain the difference in the rate of mutations between AK and melanoma lesions, as both lesions are exposed to the same environmental stress. However, the disparity in the mutation burden may be due to the difference in the cells of origin and/or the affected cellular pathways. South et al. (2014) suggested that the high mutation burden within lesions originating from keratinocytes is either because keratinocytes are more resistant to transformation, or harbour a powerful tumour suppressor activity that negatively affects the clonal expansion of the mutant tumour initiating cells.

Comparison of the frequently mutated genes observed in AKs from Chitsazzadeh et al. (2016) with those in the AKs within the current study revealed a high similarity in the more frequently mutated genes with some genes such as *FAM135B, NYAP2* and *NF1* that were mutated in 2, 4 and 2 AKs respectively in the Chitsazzadeh et al. study (2016) were not mutated in the WES AK data of the current study.

Analysis of the genetic mutations within AKs revealed a high similarity in the type of mutations between AK and cSCC lesions. In parallel to cSCC, most of the changes within AKs were missense alterations, and the majority represented UVR signature mutations in the form of C > T and CC > TT changes. Moreover, AKs contained mutations within *TP53, CDKN2A,* and members of the *RAS* and *NOTCH* gene families that have also seen to be mutated within cSCC (Brash et al., 1996, Caldeira et al., 2004, Evans et al., 2000, Forbes et

al., 2011, Kim et al., 2002, Kubo et al., 1994, Nelson et al., 1994, Nicolas et al., 2003, Oberholzer et al., 2012, Pacifico et al., 2008, Pierceall et al., 1992, Van Kranen et al., 1995, Ziegler et al., 1994). Indeed, 43 genes that were seen to be mutated in ≥ 2 of the 5 AK lesions in the WES study (i.e. ≥ 40%) were also reported to be mutated in ≥ 40% of cSCCs in the South et al. (2014). These findings suggest that in spite of the relatively benign clinical behaviour of AKs, minimal additional genetic events may be required for AKs to progress to cSCCs. On the other hand, despite the high number of mutations in AKs, these lesions can regress in 26% - 74% of cases (Frost et al., 2000, Marks et al., 1986, Thompson et al., 1993).

Whereas a number of different studies have been performed to determine the rate of progression of AKs to cSCCs (Czarnecki et al., 2002, Fuchs and Marmur, 2007, Marks et al., 1988, Mittelbronn et al., 1998), there are limited studies investigating AK progression to BCC. However, Criscione et al., (2009) showed that 36% of all primary BCCs diagnosed in their study had developed from lesions which had been previously clinically diagnosed as AKs. The two recent BCC exome sequencing studies (Bonilla et al., 2016, Jayaraman et al., 2014) identified mutations within Hedgehog signalling pathway genes (*PTCH1, SMO* and *SUFU*) with *PTCH1* mutation being highest (73% and 75%, respectively). Interestingly, *PTCH1* mutations were also observed in AKs in this thesis (in both WES and targeted sequencing). Moreover, BCCs were also observed to harbour additional mutations in other cancer-related genes that have also been seen to be mutated in cSCCs and in AKs in this thesis, including *TP53, NOTCH1* and *NOTCH2* (Bonilla et al., 2016, Jayaraman et al., 2014), while *NRAS, KRAS* or *HRAS* mutations were much lower in frequency, occurring in 2% of BCCs (Bonilla et al., 2016). While commonality of mutations in this small group of genes between AKs and BCCs does not allow one to conclude that AKs can develop into BCCs, it does lend some support to the idea that some AKs may progress to BCC.

Targeted sequencing of the selected genes provided a higher depth of coverage than that obtained in the WES data, and was conducted on a much larger sample size of AKs, as well as on PIPs, BDs and cSCCs. In the case of the later precancerous lesions (AKs and BDs) and cSCCs, the results showed a high mutation burden and UVR signature mutation as the main changes. In addition, several genes were frequently mutated in all of these lesions including *NOTCH1, TP53, MLL2, HMNC1* and *GPR98*. Admittedly, this would have

expected in cSCCs, because the genes in the targeted sequencing panel were originally chosen because they were observed to be frequently mutated in the published literature on cSCCs. The mutation frequencies between the adjacent lesions (AK/cSCC and BD/cSCC) that were analysed in chapter 4 varied, with additional mutations in the cSCC on some occasions, but in other cases an absence of mutations in the cSCC that were present in the AK or BD.

These similar mutation frequencies meant that there was no obvious additional mutation detected in the majority of cSCCs above those in the adjacent AKs and BDs, which means that this comparison did not allow the identification of a gene which frequently became mutated during the transformation of a late precancerous lesion into a cSCC. However, the benefit of the targeted sequencing in this larger group of AKs and BDs meant that this is the first study to investigate multiple genes in a moderately sized group of these late precancerous lesions in human skin (i.e. N=24 AKs and N=24 BDs, equal to 48 lesions in total with ≥ 100 read depth). The results of this study therefore provide useful information on the frequency of mutations in the selected genes in this group of lesions.

The traditional model of cancer proposes that cancer cells gain the ability of uncontrolled growth as a result of accumulation of different chronological mutations over time – called the clonal selection model (Fearon and Vogelstein, 1990, Greaves and Maley, 2012, Yates and Campbell, 2012). A recent study by Sottoriva and colleagues documenting subclonal alterations in individual glands (<10,000 cells per gland) within established colorectal cancers suggested that several/multiple driver mutations might occur early during tumorigenesis, which they termed the “Big Bang” model of cancer (Sottoriva et al., 2015). Due to subclonal growth of cells during subsequent development of the colorectal cancer, many genetic alterations are unlikely to be present through the entire tumour.

In case of skin cancer, the Big Bang model would predict that the early genetic changes would be expected to occur before the appearance of any clinically visible lesions. In a recent study by Martincorena and colleagues, the mutation burden and frequency for a list of 74 genes cancer related-genes were studied within samples from normal sun exposed skin (Martincorena et al., 2015). They identified multiple somatic mutations in epidermal cells, with ~140 mutations per cm^2 and average burden of mutation equal to 2 to 6 mutations/megabase. Similarly to cSCCs (Durinck et al., 2011, Li et al., 2015, Pickering

et al., 2014, South et al., 2014) and AKs (Chitsazzadeh et al., 2016), *TP53, NOTCH1* and *NOTCH2* were frequently mutated in the normal sun exposed skin (Martincorena et al., 2015).

It is known that skin chronically exposed to sunlight contains numerous PIPs that harbour *TP53* gene mutations in 29 - 64% of these PIPs (Jonason et al., 1996, Kramata et al., 2005, Tabata et al., 1999). Based on the fact that both cSCC and PIPs arise on chronically sun exposed skin and both conditions harbour mutations in *TP53,* some of which are shared between the two lesions, it has been hypothesised that these "patches" are early precursors of cSCC (Kramata et al., 2005, Rebel et al., 2005, Robinson et al., 2010). The fact that normal sun exposed skin in the Martincorena et al. (2015) and Chitsazzadeh et al. (2016) studies was used macroscopically (i.e. not stained or microdissected) means that many of the genetic alterations in the skin might have been either scattered randomly throughout the epidermis or, alternatively, in PIPs.

The results from the sequencing of PIPs in this thesis showed that PIPs contain mutations in a number of genes over those already known in *TP53*. Although the total number of mutations identified within PIPs (69 mutations) was less than that detected within the other lesions (AK, BD and cSCC), this finding is consistent with the classification of PIPs as an early step in the process of cSCC malignancy. In addition, the observation that all of the genes with non-silent mutations in PIPs were also mutated in the AKs, BDs and cSCCs would be compatible with a view that AKs, BDs and cSCCs may each arise from PIPs; in the case of cSCCs this might be directly from PIPs or from AKs and/or BDs which had arisen from PIPs. In addition, by comparing the genetic mutations in the lesions in this study with the published literature, the observation that many genes mutated in PIPs have also been reported as mutated in AKs, BDs and cSCCs in the literature is consistent with previous clinical studies documenting cSCCs arising from AKs and BDs (Cox et al., 1999, Criscione et al., 2009, Czarnecki et al., 2002, Fuchs and Marmur, 2007, Jaeger et al., 1999, Kao, 1986, Marks et al., 1988, Mittelbronn et al., 1998, Peterka et al., 1961). However, it is recognised that only a small number of PIPs progress to cSCCs (Rebel et al., 2005) and that it is possible that other invisible clonal lesions in human skin may also act as precursors for AKs, BDs and/or cSCCs.

Consistent with previous reports (Jonason et al., 1996, Kramata et al., 2005), the *TP53* gene was not mutated in all PIPs. Theoretically, mutations can be missed if the targeted sequenced DNA is contaminated by normal cells (e.g. keratinocytes), but, the use of target enriched sequencing in this study with careful dissection of the targeted tissue, and the fact that other genes were mutated in individual PIPs which lacked *TP53* mutations, confirm that the coding sequence of *TP53* is not mutated in a number of PIPs. It was notable that no *RAS* mutations were seen in PIPs, but as they are present in low frequency in AKs, BDs and cSCCs, it is unclear whether they are a later event in the human skin carcinogenesis pathway or simply at a low frequency in all stages of cancer development along this pathway.

The clonal mutations within *GPR98, MLL2, CDKN2A* and *CACNA1C* in PIPs indicate their occurrence early in the growth of the relevant PIPs. The subclonal nature of the other mutations in the PIPs suggests that these mutations arose later during its growth (possibly from additional UVR exposure or potentially from errors during DNA replication as a result of mutations in genes involved in cell cycle / division). Importantly, the clonality data demonstrating the presence of subclonal evolution in PIPs provides support for the "Big Bang" model of carcinogenesis (Sottoriva et al., 2015), because these groups of epidermal cells (many <3,000 cells in size) are at an early stage along the cancer development pathway, even if many PIPs do not progress further along the pathway to cSCC or BCC. There is some evidence that PIPs can regress in the absence of continued UVR exposure in mice (Remenyik et al., 2003) and sun avoidance (e.g. with the use of sunscreen) has been linked to regression of AKs in humans (Thompson et al., 1993). Thus, it appears likely that potential precancerous skin lesions such as PIPs and AKs, despite containing multiple genetic alterations, may be in a state of flux between regression and further progression along the cancer development pathway.

Overall, the work in this thesis establishes that mutations in multiple genes relevant to cancer development are present within potential precancerous skin conditions, including clinically visible (AK and BD) and clinically invisible lesions (PIPs) (figure 6.1). The results in the thesis increase the knowledge of the mutational landscape of chronically exposed skin and precancerous lesions in this organ and thus the understanding of cancer development during the precancerous stages of keratinocyte-derived skin cancer.

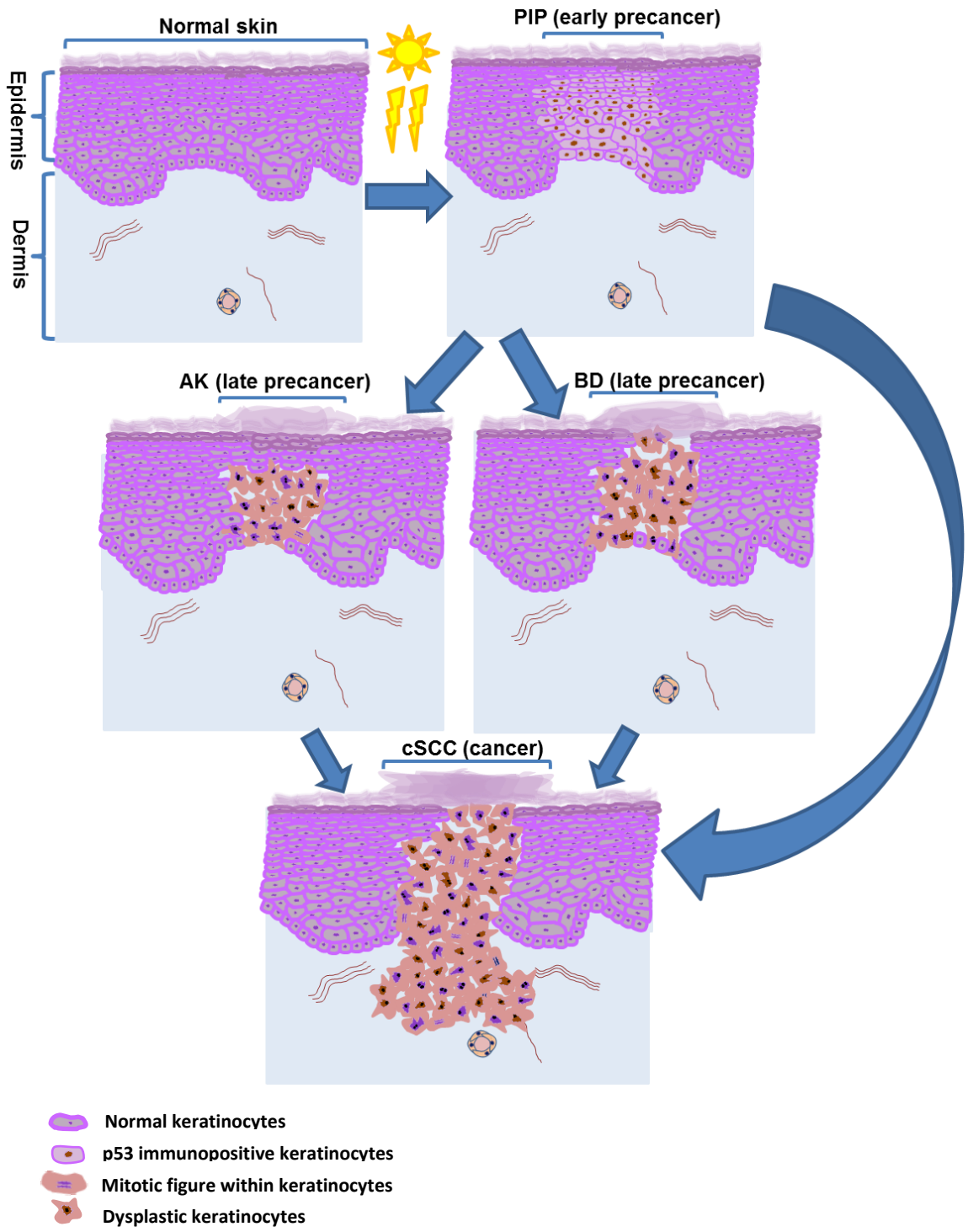


Figure 6.1 Schematic representation of skin carcinogenesis in humans. p53 immunopositive patches (PIPs, which are early precancerous lesions, frequently <3,000 cells, that develop in chronically sun-exposed skin) contain mutations in multiple genes relevant to cSCC development. Although some mutations are clonal in PIPs, many mutations are subclonal at this stage, thus providing strong support for the "Big Bang" model of cancer development. In later stage, despite the high mutation burden within the precancerous lesions such as AKs, and the fact that some may progress to cutaneous squamous cell cancer (cSCC), many of these lesions behave in a benign clinical manner and may regress spontaneously.

6.2 Challenges

Although most of the challenges have been mentioned previously in the relevant chapters, the main challenges in this project can be summarised as follows;

1. All AKs, BDs and cSCCs that were selected for WES or targeted sequencing analysis were FFPE samples. Moreover, only dysplastic cells from the AK and BD lesions and cancerous cells from cSCCs were included in the analysis. Thus obtaining DNA of good quality and sufficient quantity for analysis was challenging, but was overcome by use of multiple tissue sections from each lesion (including microdissection of up to 55 of 10µm thick sections per sample using LCM for some AKs).

2. Identification of PIPs required formalin fixation of epidermis and microdissection of the PIPs from the surrounding epidermis. It was considered that obtaining sufficient amounts of DNA from PIPs for WES was not possible. Despite detecting PIPs by performing staining on epidermal samples from 126 subjects, only 21 PIPs gave sufficient amount of DNA (≥ 100 ng) for the target enriched sequencing. Although it would have been possible to increase the DNA amount from the PIPs using whole genomic amplification, this technique was not performed due to the risk of introducing mutations reducing reliability of the results.

3. Other challenges that were also present during this thesis are the issues of money and time, something that most studies face. For example, next generation sequencing is expensive and isolation of sufficiently pure samples of DNA from many of the lesions in this thesis took significant amounts of time. This meant that while it would have been preferable to undertake WES and targeted sequencing on a greater number of lesions in each of the categories, it was not financially possible to do so and there would have been a risk that the thesis would not have been completed within the appropriate timeframe. However, these challenges allowed me to develop experience in planning research while taking funding and time-management into account.

6.3 Future directions

This thesis has opened up the possibilities for further research and future experiments to address the findings from the current project include:

1. Do PIPs contain mutations in other genes that are mutated in cSCC and/or AKs? To address this, WES on PIPs would need to be conducted.

2. PIPs less than 0.14 mm^2 in size were not sequenced in this thesis, therefore, do smaller PIPs (< 0.14 mm^2) have similar/different gene mutations? This question might be addressed in the future if lower amounts of DNA are sufficient for WES and/or targeted sequencing as technological improvements in next generation sequencing are made. Furthermore, the advent of single-cell sequencing methodologies should allow the determination of the mutational landscape of (apparently) normal keratinocytes between sun-exposed and non-sun-exposed skin. This could provide insight into the frequency of mutation induced by UVR exposure.

3. In addition to studying the coding mutations within PIPs, other molecular studies including examining promoter methylation status might be helpful.

4. Functional analysis of genes mutated in PIPs, AKs, BDs and cSCCs using *in vivo* (mice models) or *in vitro* (primary keratinocytes or HaCaT cells) assays by expressing wild-type/mutant alleles in these systems might provide information on their role as tumour suppressor genes or oncogenes within keratinocytes during the early stages of cancer development. This type of analysis might be particularly useful for genes such as *GPR98* and *MLL2* where there is little information about their role in keratinocytes / skin.

References

Ackerman, A. B. & Mones, J. M. (2006). Solar (actinic) keratosis is squamous cell carcinoma. *Br J Dermatol,* 155,(1)**,** 9-22.

Acosta, H., Lopez, S. L., Revinski, D. R. & Carrasco, A. E. (2011). Notch destabilises maternal beta-catenin and restricts dorsal-anterior development in Xenopus. *Development,* 138,(12)**,** 2567-79.

Agilent Technologies. (2014). *SureSelectXT Target Enrichment System for Illumina Paired-End Multiplexed Sequencing Library.* [Online]. USA. Available: http://www.agilent.com/cs/library/usermanuals/Public/G7530-90000.pdf. [Accessed 27 April 2016].

Ahmadian, A., Ren, Z. P., Williams, C., Ponten, F., Odeberg, J., Ponten, J., Uhlen, M. & Lundeberg, J. (1998). Genetic instability in the 9q22.3 region is a late event in the development of squamous cell carcinoma. *Oncogene,* 17,(14)**,** 1837-43.

Akbani, R., Akdemir, K. C., Aksoy, B. A., Albert, M., Ally, A., Amin, S. B., Arachchi, H., Arora, A., Auman, J. T., Ayala, B., Baboud, J., Balasundaram, M., Balu, S., Barnabas, N., Bartlett, J., Bartlett, P., Bastian, B. C., Baylin, S. B., Behera, M., Belyaev, D., et al. (2015). Genomic classification of cutaneous melanoma. *Cell,* 161,(7)**,** 1681-96.

Alam, M. & Ratner, D. (2001). Cutaneous squamous-cell carcinoma. *N Engl J Med,* 344,(13)**,** 975-83.

Aleem, E., Kiyokawa, H. & Kaldis, P. (2005). Cdc2-cyclin E complexes regulate the G1/S phase transition. *Nat Cell Biol,* 7,(8)**,** 831-6.

Amaral, J. D., Xavier, J. M., Steer, C. J. & Rodrigues, C. M. (2010). Targeting the p53 pathway of apoptosis. *Curr Pharm Des,* 16,(22)**,** 2493-503.

Amsen, D., Antov, A. & Flavell, R. A. (2009). The different faces of Notch in T-helper-cell differentiation. *Nat Rev Immunol,* 9,(2)**,** 116-24.

Antzelevitch, C., Pollevick, G. D., Cordeiro, J. M., Casis, O., Sanguinetti, M. C., Aizawa, Y., Guerchicoff, A., Pfeiffer, R., Oliva, A., Wollnik, B., Gelber, P., Bonaros, E. P., Burashnikov, E., Wu, Y., Sargent, J. D., Schickel, S., Oberheiden, R., Bhatia, A., Hsu, L. F., Haissaguerre, M., et al. (2007). Loss-of-function mutations in the cardiac calcium channel underlie a new clinical entity characterized by ST-segment elevation, short QT intervals, and sudden cardiac death. *Circulation,* 115,(4)**,** 442-9.

Applied Biosystems: Hitachi. (2010). *Applied Biosystems 3730/3730xl DNA Analyzers: User Guide.* [Online]. Available: http://tools.thermofisher.com/content/sfs/manuals/cms_041259.pdf. [Accessed 23 June 2016].

Arlette, J. P. & Trotter, M. J. (2004). Squamous cell carcinoma in situ of the skin: history, presentation, biology and treatment. *Australas J Dermatol,* 45,(1)**,** 1-9.

Artavanis-Tsakonas, S., Rand, M. D. & Lake, R. J. (1999). Notch signaling: cell fate control and signal integration in development. *Science,* 284,(5415)**,** 770-6.

Asgari, M. M., Wang, W., Ioannidis, N. M., Itnyre, J., Hoffmann, T., Jorgenson, E. & Whittemore, A. S. (2016). Identification of susceptibility loci for cutaneous squamous cell carcinoma. *J Invest Dermatol,* 136,(5)**,** 930-7.

Aston, K. I., Krausz, C., Laface, I., Ruiz-Castane, E. & Carrell, D. T. (2010). Evaluation of 172 candidate polymorphisms for association with oligozoospermia or azoospermia in a large cohort of men of European descent. *Hum Reprod,* 25,(6)**,** 1383-97.

Ayyanan, A., Civenni, G., Ciarloni, L., Morel, C., Mueller, N., Lefort, K., Mandinova, A., Raffoul, W., Fiche, M., Dotto, G. P. & Brisken, C. (2006). Increased Wnt signaling triggers oncogenic conversion of human breast epithelial cells by a Notch-dependent mechanism. *Proc Natl Acad Sci U S A,* 103,(10)**,** 3799-804.

Backvall, H., Stromberg, S., Gustafsson, A., Asplund, A., Sivertsson, A., Lundeberg, J. & Ponten, F. (2004a). Mutation spectra of epidermal p53 clones adjacent to basal cell carcinoma and squamous cell carcinoma. *Exp Dermatol,* 13,(10)**,** 643-50.

Backvall, H., Wolf, O., Hermelin, H., Weitzberg, E. & Ponten, F. (2004b). The density of epidermal p53 clones is higher adjacent to squamous cell carcinoma in comparison with basal cell carcinoma. *Br J Dermatol,* 150,(2)**,** 259-66.

Bagazgoitia, L., Cuevas, J. & Juarranz, A. (2010). Expression of p53 and p16 in actinic keratosis, bowenoid actinic keratosis and Bowen's disease. *J Eur Acad Dermatol Venereol,* 24,(2)**,** 228-30.

Balint, K., Xiao, M., Pinnix, C. C., Soma, A., Veres, I., Juhasz, I., Brown, E. J., Capobianco, A. J., Herlyn, M. & Liu, Z. J. (2005). Activation of Notch1 signaling is required for beta-catenin-mediated human primary melanoma progression. *J Clin Invest,* 115,(11)**,** 3166-76.

Bastiaens, M. T., Ter Huurne, J. A., Kielich, C., Gruis, N. A., Westendorp, R. G., Vermeer, B. J. & Bavinck, J. N. (2001). Melanocortin-1 receptor gene variants determine the risk of nonmelanoma skin cancer independently of fair skin and red hair. *Am J Hum Genet,* 68,(4)**,** 884-94.

Beaumont, K. A., Smit, D. J., Liu, Y. Y., Chai, E., Patel, M. P., Millhauser, G. L., Smith, J. J., Alewood, P. F. & Sturm, R. A. (2012). Melanocortin-1 receptor-mediated signalling pathways activated by NDP-MSH and HBD3 ligands. *Pigment Cell Melanoma Res,* 25,(3)**,** 370-4.

Behrens, J. (2000). Control of beta-catenin signaling in tumor development. *Ann N Y Acad Sci,* 910**,** 21-33.

Behrens, J., Jerchow, B. A., Wurtele, M., Grimm, J., Asbrand, C., Wirtz, R., Kuhl, M., Wedlich, D. & Birchmeier, W. (1998). Functional interaction of an axin homolog, conductin, with beta-catenin, APC, and GSK3beta. *Science,* 280,(5363)**,** 596-9.

Beissert, S., Hosoi, J., Kuhn, R., Rajewsky, K., Muller, W. & Granstein, R. D. (1996). Impaired immunosuppressive response to ultraviolet radiation in interleukin-10-deficient mice. *J Invest Dermatol,* 107,(4)**,** 553-7.

Benjamin, C. L. & Ananthaswamy, H. N. (2007). p53 and the pathogenesis of skin cancer. *Toxicol Appl Pharmacol,* 224,(3)**,** 241-8.

Bennett, M. A., O'Grady, A., Kay, E. W., Leader, M. & Murphy, G. M. (1997). p53 mutations in squamous cell carcinomas from renal transplant recipients. *Biochem Soc Trans,* 25,(1)**,** 342-5.

Berg, R. J., Van Kranen, H. J., Rebel, H. G., De Vries, A., Van Vloten, W. A., Van Kreijl, C. F., Van Der Leun, J. C. & De Gruijl, F. R. (1996). Early p53 alterations in mouse skin carcinogenesis by UVB radiation: immunohistochemical detection of mutant p53 protein in clusters of preneoplastic epidermal cells. *Proc Natl Acad Sci U S A,* 93,(1)**,** 274-8.

Berger, M. F., Hodis, E., Heffernan, T. P., Deribe, Y. L., Lawrence, M. S., Protopopov, A., Ivanova, E., Watson, I. R., Nickerson, E., Ghosh, P., Zhang, H., Zeid, R., Ren, X., Cibulskis, K., Sivachenko, A. Y., Wagle, N., Sucker, A., Sougnez, C., Onofrio, R., Ambrogio, L., et al. (2012). Melanoma genome sequencing reveals frequent PREX2 mutations. *Nature,* 485,(7399)**,** 502-6.

Berglund, L., Bjorling, E., Oksvold, P., Fagerberg, L., Asplund, A., Szigyarto, C. A., Persson, A., Ottosson, J., Wernerus, H., Nilsson, P., Lundberg, E., Sivertsson, A., Navani, S., Wester, K., Kampf, C., Hober, S., Ponten, F. & Uhlen, M. (2008). A genecentric Human Protein Atlas for expression profiles based on antibodies. *Mol Cell Proteomics,* 7,(10)**,** 2019-27.

Berlin, J. M. (2010). Current and emerging treatment strategies for the treatment of actinic keratosis. *Clin Cosmet Investig Dermatol,* 3**,** 119-26.

Berman, B. & Cockerell, C. J. (2013). Pathobiology of actinic keratosis: ultraviolet-dependent keratinocyte proliferation. *J Am Acad Dermatol* 68,(Suppl 1)**,** S10-9.

Bhagat, R., Kumar, S. S., Vaderhobli, S., Premalata, C. S., Pallavi, V. R., Ramesh, G. & Krishnamoorthy, L. (2014). Epigenetic alteration of p16 and retinoic acid receptor beta genes in the development of epithelial ovarian carcinoma. *Tumour Biol,* 35,(9)**,** 9069-78.

Bhatia, N. & Spiegelman, V. S. (2005). Activation of Wnt/beta-catenin/Tcf signaling in mouse skin carcinogenesis. *Mol Carcinog,* 42,(4)**,** 213-21.

Blaumueller, C. M., Qi, H., Zagouras, P. & Artavanis-Tsakonas, S. (1997). Intracellular cleavage of Notch leads to a heterodimeric receptor on the plasma membrane. *Cell,* 90,(2)**,** 281-91.

Bode, A. M. & Dong, Z. (2004). Post-translational modification of p53 in tumorigenesis. *Nat Rev Cancer,* 4,(10)**,** 793-805.

Bohm, M., Wolff, I., Scholzen, T. E., Robinson, S. J., Healy, E., Luger, T. A., Schwarz, T. & Schwarz, A. (2005). alpha-Melanocyte-stimulating hormone protects from ultraviolet radiation-induced apoptosis and DNA damage. *J Biol Chem,* 280,(7)**,** 5795-802.

Boldrini, L., Loggini, B., Gisfredi, S., Zucconi, Y., Baldinotti, F., Fogli, A., Simi, P., Cervadoro, G., Barachini, P., Basolo, F., Pingitore, R. & Fontanini, G. (2003). Mutations of Fas (APO-1/CD95) and p53 genes in nonmelanoma skin cancer. *J Cutan Med Surg,* 7,(2)**,** 112-8.

Bolshakov, S., Walker, C. M., Strom, S. S., Selvan, M. S., Clayman, G. L., El-Naggar, A., Lippman, S. M., Kripke, M. L. & Ananthaswamy, H. N. (2003). p53 mutations in human aggressive and nonaggressive basal and squamous cell carcinomas. *Clin Cancer Res,* 9,(1)**,** 228-34.

Bonilla, X., Parmentier, L., King, B., Bezrukov, F., Kaya, G., Zoete, V., Seplyarskiy, V. B., Sharpe, H. J., McKee, T., Letourneau, A., Ribaux, P. G., Popadin, K., Basset-Seguin, N., Chaabene, R. B., Santoni, F. A., Andrianova, M. A., Guipponi, M., Garieri, M., Verdan, C., Grosdemange, K., et al. (2016). Genomic analysis identifies new drivers and progression pathways in skin basal cell carcinoma. *Nat Genet,* 48,(4)**,** 398-406.

Bos, J. L. (1989). Ras oncogenes in human cancer: a review. *Cancer Res,* 49,(17)**,** 4682-9.

Boukamp, P. (2005). Non-melanoma skin cancer: what drives tumor development and progression? *Carcinogenesis,* 26,(10)**,** 1657-67.

Bouwes-Bavinck, J. N., Hardie, D. R., Green, A., Cutmore, S., MacNaught, A., O'Sullivan, B., Siskind, V., Van Der Woude, F. J. & Hardie, I. R. (1996). The risk of skin cancer in renal transplant recipients in Queensland, Australia. A follow-up study. *Transplantation,* 61,(5)**,** 715-21.

Box, N. F., Duffy, D. L., Irving, R. E., Russell, A., Chen, W., Griffyths, L. R., Parsons, P. G., Green, A. C. & Sturm, R. A. (2001). Melanocortin-1 receptor genotype is a risk factor for basal and squamous cell carcinoma. *J Invest Dermatol,* 116,(2)**,** 224-9.

Boyd, A. S., Stasko, T., Cameron, G. S., Russell, M. & King, L. E. (2001). Histologic features of actinic keratoses in solid organ transplant recipients and healthy controls. *J Am Acad Dermatol* 45,(2)**,** 217-21.

Brantsch, K. D., Meisner, C., Schönfisch, B., Trilling, B., Wehner-Caroli, J., Röcken, M. & Breuninger, H. (2008). Analysis of risk factors determining prognosis of cutaneous squamous-cell carcinoma: a prospective study. *Lancet Oncol,* 9,(8)**,** 713-720.

Brasanac, D., Boricic, I., Todorovic, V., Tomanovic, N. & Radojevic, S. (2005). Cyclin A and beta-catenin expression in actinic keratosis, Bowen's disease and invasive squamous cell carcinoma of the skin. *Br J Dermatol,* 153,(6)**,** 1166-75.

Brash, D. E. (1988). UV mutagenic photoproducts in Escherichia coli and human cells: a molecular genetics perspective on human skin cancer. *Photochem Photobiol,* 48,(1)**,** 59-66.

Brash, D. E. (2006). Roles of the transcription factor p53 in keratinocyte carcinomas. *Br J Dermatol,* 154 (Suppl 1)**,** S8-10.

Brash, D. E. (2015). UV signature mutations. *Photochem Photobiol,* 91,(1)**,** 15-26.

Brash, D. E., Rudolph, J. A., Simon, J. A., Lin, A., McKenna, G. J., Baden, H. P., Halperin, A. J. & Ponten, J. (1991). A role for sunlight in skin cancer: UV-induced p53 mutations in squamous cell carcinoma. *Proc Natl Acad Sci U S A,* 88,(22)**,** 10124-8.

Brash, D. E., Ziegler, A., Jonason, A. S., Simon, J. A., Kunala, S. & Leffell, D. J. (1996). Sunlight and sunburn in human skin cancer: p53, apoptosis, and tumor promotion. *J Investig Dermatol Symp Proc,* 1,(2)**,** 136-42.

Brou, C., Logeat, F., Gupta, N., Bessia, C., LeBail, O., Doedens, J. R., Cumano, A., Roux, P., Black, R. A. & Israel, A. (2000). A novel proteolytic cleavage involved in Notch signaling: the role of the disintegrin-metalloprotease TACE. *Mol Cell,* 5,(2)**,** 207-16.

Brown, V. L., Harwood, C. A., Crook, T., Cronin, J. G., Kelsell, D. P. & Proby, C. M. (2004). p16INK4a and p14ARF tumor suppressor genes are commonly inactivated in cutaneous squamous cell carcinoma. *J Invest Dermatol,* 122,(5)**,** 1284-92.

Burns, J. E., Baird, M. C., Clark, L. J., Burns, P. A., Edington, K., Chapman, C., Mitchell, R., Robertson, G., Soutar, D. & Parkinson, E. K. (1993). Gene mutations and increased levels of p53 protein in human squamous cell carcinomas and their cell lines. *Br J Cancer,* 67,(6)**,** 1274-84.

Caldeira, S., Filotico, R., Accardi, R., Zehbe, I., Franceschi, S. & Tommasino, M. (2004). p53 mutations are common in human papillomavirus type 38-positive non-melanoma skin cancers. *Cancer Lett,* 209,(1)**,** 119-24.

Callen, J. P., Bickers, D. R. & Moy, R. L. (1997). Actinic keratoses. *J Am Acad Dermatol,* 36,(4)**,** 650-3.

Campbell, C., Quinn, A. G. & Rees, J. L. (1993a). Codon 12 Harvey-ras mutations are rare events in non-melanoma human skin cancer. *Br J Dermatol,* 128,(2)**,** 111-4.

Campbell, C., Quinn, A. G., Ro, Y. S., Angus, B. & Rees, J. L. (1993b). p53 mutations are common and early events that precede tumor invasion in squamous cell neoplasia of the skin. *J Invest Dermatol,* 100,(6)**,** 746-8.

Capaccione, K. M. & Pine, S. R. (2013). The Notch signaling pathway as a mediator of tumor survival. *Carcinogenesis,* 34,(7)**,** 1420-30.

Carter, S. L., Cibulskis, K., Helman, E., McKenna, A., Shen, H., Zack, T., Laird, P. W., Onofrio, R. C., Winckler, W., Weir, B. A., Beroukhim, R., Pellman, D., Levine, D. A., Lander, E. S., Meyerson, M. & Getz, G. (2012). Absolute quantification of somatic DNA alterations in human cancer. *Nat Biotech,* 30,(5)**,** 413-421.

Cassarino, D. S., Derienzo, D. P. & Barr, R. J. (2006). Cutaneous squamous cell carcinoma: a comprehensive clinicopathologic classification--part two. *J Cutan Pathol,* 33,(4)**,** 261-79.

Castano, E., Rodriguez-Peralto, J. L., Lopez-Rios, F., Gomez, C., Zimmermann, M. & Iglesias-Diez, L. (2002). Keratinocyte dysplasia: an usual finding after transplantation or chemotherapy. *J Cutan Pathol,* 29,(10)**,** 579-84.

Caulin, C., Nguyen, T., Lang, G. A., Goepfert, T. M., Brinkley, B. R., Cai, W. W., Lozano, G. & Roop, D. R. (2007). An inducible mouse model for skin cancer reveals distinct roles for gain- and loss-of-function p53 mutations. *J Clin Invest,* 117,(7)**,** 1893-901.

Chahal, H. S., Lin, Y., Ransohoff, K. J., Hinds, D. A., Wu, W., Dai, H. J., Qureshi, A. A., Li, W. Q., Kraft, P., Tang, J. Y., Han, J. & Sarin, K. Y. (2016). Genome-wide association study identifies novel susceptibility loci for cutaneous squamous cell carcinoma. *Nat Commun,* 7**,** 12048-56.

Chang, E. H., Furth, M. E., Scolnick, E. M. & Lowy, D. R. (1982a). Tumorigenic transformation of mammalian cells induced by a normal human gene homologous to the oncogene of Harvey murine sarcoma virus. *Nature,* 297,(5866)**,** 479-83.

Chang, E. H., Gonda, M. A., Ellis, R. W., Scolnick, E. M. & Lowy, D. R. (1982b). Human genome contains four genes homologous to transforming genes of Harvey and Kirsten murine sarcoma viruses. *Proc Natl Acad Sci U S A,* 79,(16)**,** 4848-52.

Chen, Y., Li, D., Liu, H., Xu, H., Zheng, H., Qian, F., Li, W., Zhao, C., Wang, Z. & Wang, X. (2011). Notch-1 signaling facilitates survivin expression in human non-small cell lung cancer cells. *Cancer Biol Ther,* 11,(1)**,** 14-21.

Cherpelis, B. S., Marcusen, C. & Lang, P. G. (2002). Prognostic factors for metastasis in squamous cell carcinoma of the skin. *Dermatol Surg,* 28,(3)**,** 268-73.

Chikayama, S., Sugano, T., Takahashi, Y., Ikeda, M., Kimura, S., Kobayashi, Y. & Kondo, M. (2000). Nuclear accumulation of p53 protein and apoptosis induced by various anticancer agents, u.v.-irradiation and heat shock in primary normal human skin fibroblasts. *Int J Oncol,* 16,(6)**,** 1117-24.

Chiles, M. C., Ai, L., Zuo, C., Fan, C. Y. & Smoller, B. R. (2003). E-cadherin promoter hypermethylation in preneoplastic and neoplastic skin lesions. *Mod Pathol,* 16,(10)**,** 1014-8.

Chitsazzadeh, V., Coarfa, C., Drummond, J. A., Nguyen, T., Joseph, A., Chilukuri, S., Charpiot, E., Adelmann, C. H., Ching, G., Nguyen, T. N., Nicholas, C., Thomas, V. D., Migden, M., MacFarlane, D., Thompson, E., Shen, J., Takata, Y., McNiece, K., Polansky, M. A., Abbas, H. A., et al. (2016). Cross-species identification of genomic drivers of squamous cell carcinoma development across preneoplastic intermediates. *Nat Commun,* 7**,** 12601-18.

Chitsazzadeh, V., Xiao, W., Gunaratne, P., Coarfa, C., Su, X., Nguyen, T. H., Thomas, V. D., Joseph, A. K., Flores, E. R. & Tsai, K. Y.(2013). Abstract 815: Integrated genomic analysis of cutaneous squamous cell carcinoma progression. IN: *Proceedings of the 104th Annual Meeting of the American Association for Cancer Research*, Washington, DC. Philadelphia (PA). 15 April 2013. *Cancer Res*, 815.

Cho, Y., Gorina, S., Jeffrey, P. D. & Pavletich, N. P. (1994). Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science,* 265,(5170)**,** 346-55.

Christensen, S. R., McNiff, J. M., Cool, A. J., Aasi, S. Z., Hanlon, A. M. & Leffell, D. J. (2016). Histopathologic assessment of depth of follicular invasion of squamous cell carcinoma (SCC) in situ (SCCis): Implications for treatment approach. *J Am Acad Dermatol,* 74,(2)**,** 356-62.

Chute, C. G., Chuang, T. Y., Bergstralh, E. J. & Su, W. P. (1991). The subsequent risk of internal cancer with Bowen's disease. A population-based study. *JAMA,* 266,(6)**,** 816-9.

Cockerell, C. J. (2000). Histopathology of incipient intraepidermal squamous cell carcinoma ("actinic keratosis"). *J Am Acad Dermatol,* 42,(Supp 1)**,** S11-17.

Cockerell, C. J. & Wharton, J. R. (2005). New histopathological classification of actinic keratosis (incipient intraepidermal squamous cell carcinoma). *J Drugs Dermatol,* 4,(4)**,** 462-7.

Cox, A. D. & Der, C. J. (2010). Ras history: The saga continues. *Small GTPases,* 1,(1)**,** 2-27.

Cox, N. H., Eedy, D. J. & Morton, C. A. (1999). Guidelines for management of Bowen's disease. British Association of Dermatologists. *Br J Dermatol,* 141,(4)**,** 633-41.

Cox, N. H., Eedy, D. J. & Morton, C. A. (2007). Guidelines for management of Bowen's disease: 2006 update. *Br J Dermatol,* 156,(1)**,** 11-21.

Criscione, V. D., Weinstock, M. A., Naylor, M. F., Luque, C., Eide, M. J. & Bingham, S. F. (2009). Actinic keratoses: Natural history and risk of malignant transformation in the Veterans Affairs Topical Tretinoin Chemoprevention Trial. *Cancer,* 115,(11)**,** 2523-30.

Cumberbatch, M., Dearman, R. J., Griffiths, C. E. & Kimber, I. (2000). Langerhans cell migration. *Clin Exp Dermatol,* 25,(5)**,** 413-8.

Cummings, M., McGinley, C. V., Wilkinson, N., Field, S. L., Duffy, S. R. & Orsi, N. M. (2011). A robust RNA integrity-preserving staining protocol for laser capture microdissection of endometrial cancer tissue. *Anal Biochem,* 416,(1)**,** 123-5.

Czarnecki, D., Meehan, C. J., Bruce, F. & Culjak, G. (2002). The majority of cutaneous squamous cell carcinomas arise in actinic keratoses. *J Cutan Med Surg,* 6,(3)**,** 207-9.

D'Orazio, J., Jarrett, S., Amaro-Ortiz, A. & Scott, T. (2013). UV radiation and the skin. *Int J Mol Sci,* 14,(6)**,** 12222-48.

Dash, B. C. & El-Deiry, W. S. (2005). Phosphorylation of p21 in G2/M promotes cyclin B-Cdc2 kinase activity. *Mol Cell Biol,* 25,(8)**,** 3364-87.

Davies, H., Bignell, G. R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M. J., Bottomley, W., Davis, N., Dicks, E., Ewing, R., Floyd, Y., Gray, K., Hall, S., Hawes, R., Hughes, J., Kosmidou, V., Menzies, A., et al. (2002). Mutations of the BRAF gene in human cancer. *Nature,* 417,(6892)**,** 949-54.

De Berker, D., McGregor, J. M. & Hughes, B. R. (2007). Guidelines for the management of actinic keratoses. *Br J Dermatol,* 156,(2)**,** 222-30.

De Graaf, Y. G., Rebel, H., Elghalbzouri, A., Cramers, P., Nellen, R. G., Willemze, R., Bouwes-Bavinck, J. N. & De Gruijl, F. R. (2008). More epidermal p53 patches adjacent to skin carcinomas in renal transplant recipients than in immunocompetent patients: the role of azathioprine. *Exp Dermatol,* 17,(4)**,** 349-55.

De Gruijl, F. R. & Rebel, H. (2008). Early events in UV carcinogenesis--DNA damage, target cells and mutant p53 foci. *Photochem Photobiol,* 84,(2)**,** 382-7.

DeLeo, A. B., Jay, G., Appella, E., Dubois, G. C., Law, L. W. & Old, L. J. (1979). Detection of a transformation-related antigen in chemically induced sarcomas and other transformed cells of the mouse. *Proc Natl Acad Sci U S A,* 76,(5)**,** 2420-4.

Demehri, S., Turkoz, A. & Kopan, R. (2009). Epidermal Notch1 loss promotes skin tumorigenesis by impacting the stromal microenvironment. *Cancer Cell,* 16,(1)**,** 55-66.

Dotto, G. P., O'Connell, J., Patskan, G., Conti, C., Ariza, A. & Slaga, T. J. (1988). Malignant progression of papilloma-derived keratinocytes: differential effects of the ras, neu, and p53 oncogenes. *Mol Carcinog,* 1,(3)**,** 171-9.

Drake, A. L. & Walling, H. W. (2008). Variations in presentation of squamous cell carcinoma in situ (Bowen's disease) in immunocompromised patients. *J Am Acad Dermatol,* 59,(1)**,** 68-71.

Dumaz, N., Van Kranen, H. J., De Vries, A., Berg, R. J., Wester, P. W., Van Kreijl, C. F., Sarasin, A., Daya-Grosjean, L. & De Gruijl, F. R. (1997). The role of UV-B light in skin carcinogenesis through the analysis of p53 mutations in squamous cell carcinomas of hairless mice. *Carcinogenesis,* 18,(5)**,** 897-904.

Dunsch, A. K., Linnane, E., Barr, F. A. & Gruneberg, U. (2011). The astrin-kinastrin/SKAP complex localizes to microtubule plus ends and facilitates chromosome alignment. *J Cell Biol,* 192,(6)**,** 959-68.

Durinck, S., Ho, C., Wang, N. J., Liao, W., Jakkula, L. R., Collisson, E. A., Pons, J., Chan, S. W., Lam, E. T., Chu, C., Park, K., Hong, S. W., Hur, J. S., Huh, N., Neuhaus, I. M., Yu, S. S., Grekin, R. C., Mauro, T. M., Cleaver, J. E., Kwok, P. Y., et al. (2011). Temporal dissection of tumorigenesis in primary cancers. *Cancer Discov,* 1,(2)**,** 137-43.

Eedy, D. J. & Gavin, A. T. (1987). Thirteen-year retrospective study of Bowen's disease in Northern Ireland. *Br J Dermatol,* 117,(6)**,** 715-20.

Einspahr, J., Alberts, D. S., Aickin, M., Welch, K., Bozzo, P., Grogan, T. & Nelson, M. (1997). Expression of p53 protein in actinic keratosis, adjacent, normal-appearing, and non-sun-exposed human skin. *Cancer Epidemiol Biomarkers Prev,* 6,(8)**,** 583-7.

Ensembl. [Online]. Available: www.ensembl.org/index.html. [Accessed 15 August 2016].

Ensembl. *Variant Effect Predictor.* [Online]. Available: www.ensembl.org/info/docs/tools/vep/. [Accessed 23 August 2016].

Euvrard, S., Kanitakis, J., Decullier, E., Butnaru, A. C., Lefrancois, N., Boissonnat, P., Sebbag, L., Garnier, J. L., Pouteil-Noble, C., Cahen, R., Morelon, E., Touraine, J. L., Claudy, A. &

Chapuis, F. (2006). Subsequent skin cancers in kidney and heart transplant recipients after the first squamous cell carcinoma. *Transplantation,* 81,(8)**,** 1093-100.

Evans, T., Boonchai, W., Shanley, S., Smyth, I., Gillies, S., Georgas, K., Wainwright, B., Chenevix-Trench, G. & Wicking, C. (2000). The spectrum of patched mutations in a collection of Australian basal cell carcinomas. *Hum Mutat,* 16,(1)**,** 43-8.

Fearon, E. R. & Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. *Cell,* 61,(5)**,** 759-67.

Feitosa, N. M., Zhang, J., Carney, T. J., Metzger, M., Korzh, V., Bloch, W. & Hammerschmidt, M. (2012). Hemicentin 2 and Fibulin 1 are required for epidermal-dermal junction formation and fin mesenchymal cell migration during zebrafish development. *Dev Biol,* 369,(2)**,** 235-48.

Fernandez-Medarde, A. & Santos, E. (2011). Ras in cancer and developmental diseases. *Genes Cancer,* 2,(3)**,** 344-58.

Findlay, G. M. (1928). Ultra-Violet light and skin cancer. *Lancet,* 212,(5491)**,** 1070-3.

Fiorentino, D. F., Zlotnik, A., Vieira, P., Mosmann, T. R., Howard, M., Moore, K. W. & O'Garra, A. (1991). IL-10 acts on the antigen-presenting cell to inhibit cytokine production by Th1 cells. *J Immunol,* 146,(10)**,** 3444-51.

Fontenot, J. D., Rasmussen, J. P., Williams, L. M., Dooley, J. L., Farr, A. G. & Rudensky, A. Y. (2005). Regulatory T cell lineage specification by the forkhead transcription factor foxp3. *Immunity,* 22,(3)**,** 329-41.

Foo, C. C., Lee, J. S., Guilanno, V., Yan, X., Tan, S. H. & Giam, Y. C. (2007). Squamous cell carcinoma and Bowen's disease of the skin in Singapore. *Ann Acad Med Singapore,* 36,(3)**,** 189-93.

Forbes, S. A., Bindal, N., Bamford, S., Cole, C., Kok, C. Y., Beare, D., Jia, M., Shepherd, R., Leung, K., Menzies, A., Teague, J. W., Campbell, P. J., Stratton, M. R. & Futreal, P. A. (2011). COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res,* 39,(Database issue)**,** D945-50.

Forbes, S. A., Tang, G., Bindal, N., Bamford, S., Dawson, E., Cole, C., Kok, C. Y., Jia, M., Ewing, R., Menzies, A., Teague, J. W., Stratton, M. R. & Futreal, P. A. (2010). COSMIC (the Catalogue of Somatic Mutations in Cancer): a resource to investigate acquired mutations in human cancer. *Nucleic Acids Res,* 38,(Database issue)**,** D652-7.

Foulds, L. (1954). The experimental study of tumor progression: a review. *Cancer Res,* 14,(5)**,** 327-39.

Fransen, M., Karahalios, A., Sharma, N., English, D. R., Giles, G. G. & Sinclair, R. D. (2012). Non-melanoma skin cancer in Australia. *Med J Aust,* 197,(10)**,** 565-8.

Friedberg, E. C. (2001). How nucleotide excision repair protects against cancer. *Nat Rev Cancer,* 1,(1)**,** 22-33.

Frost, C., Williams, G. & Green, A. (2000). High incidence and regression rates of solar keratoses in a queensland community. *J Invest Dermatol,* 115,(2)**,** 273-7.

Frost, C. A. & Green, A. C. (1994). Epidemiology of solar keratoses. *Br J Dermatol,* 131,(4)**,** 455-64.

Fuchs, A. & Marmur, E. (2007). The kinetics of skin cancer: progression of actinic keratosis to squamous cell carcinoma. *Dermatol Surg,* 33,(9)**,** 1099-101.

Glogau, R. G. (2000). The risk of progression to invasive disease. *J Am Acad Dermatol,* 42,(1)**,** 23-4.

Gostissa, M., Ranganath, S., Bianco, J. M. & Alt, F. W. (2009). Chromosomal location targets different MYC family gene members for oncogenic translocations. *Proc Natl Acad Sci U S A,* 106,(7)**,** 2265-70.

Greaves, M. & Maley, C. C. (2012). Clonal evolution in cancer. *Nature,* 481,(7381)**,** 306-13.

Green, A., Williams, G., Neale, R., Hart, V., Leslie, D., Parsons, P., Marks, G. C., Gaffney, P., Battistutta, D., Frost, C., Lang, C. & Russell, A. (1999). Daily sunscreen application and betacarotene supplementation in prevention of basal-cell and squamous-cell carcinomas of the skin: a randomised controlled trial. *Lancet,* 354,(9180)**,** 723-9.

Griewank, K. G., Murali, R., Schilling, B., Schimming, T., Moller, I., Moll, I., Schwamborn, M., Sucker, A., Zimmer, L., Schadendorf, D. & Hillen, U. (2013). TERT promoter mutations are frequent in cutaneous basal cell carcinoma and squamous cell carcinoma. *PLOS One,* 8,(11)**,** 1-5.

Gruner, S., Oesterwitz, H., Stoppe, H., Henke, W., Eckert, R. & Sonnichsen, N. (1992). Cis-urocanic acid as a mediator of ultraviolet-light-induced immunosuppression. *Semin Hematol,* 29,(2)**,** 102-7.

Guinea-Viniegra, J., Zenz, R., Scheuch, H., Jimenez, M., Bakiri, L., Petzelbauer, P. & Wagner, E. F. (2012). Differentiation-induced skin cancer suppression by FOS, p53, and TACE/ADAM17. *J Clin Invest,* 122,(8)**,** 2898-910.

Guy, G. P., Machlin, S. R., Ekwueme, D. U. & Yabroff, K. R. (2015). Prevalence and costs of skin cancer treatment in the U.S., 2002-2006 and 2007-2011. *Am J Prev Med,* 48,(2)**,** 183-7.

Haake, A., Scott, G. A. & Holbrook, K. A. (2001). Structure and function of the skin: overview of the epidermis and dermis. *In:* Woodly, D. T. & Freinkel, R. K. (eds.) IN: *The biology of the skin.* New York: The Parthoenon Publishing Group Inc.

Hama, N., Ohtsuka, T. & Yamazaki, S. (2006). Detection of mucosal human papilloma virus DNA in bowenoid papulosis, Bowen's disease and squamous cell carcinoma of the skin. *J Dermatol,* 33,(5)**,** 331-7.

Hamzehloie, T., Mojarrad, M., Hasanzadeh-Nazarabadi, M. & Shekouhi, S. (2012). The role of tumor protein 53 mutations in common human cancers and targeting the murine double minute 2-p53 interaction for cancer therapy. *Iran J Med Sci,* 37,(1)**,** 3-8.

Han, J., Kraft, P., Nan, H., Guo, Q., Chen, C., Qureshi, A., Hankinson, S. E., Hu, F. B., Duffy, D. L., Zhao, Z. Z., Martin, N. G., Montgomery, G. W., Hayward, N. K., Thomas, G., Hoover, R. N., Chanock, S. & Hunter, D. J. (2008). A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLOS Genet,* 4,(5)**,** 1-11.

Hanahan, D. & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell,* 144,(5)**,** 646-74.

Hansen, J. P., Drake, A. L. & Walling, H. W. (2008). Bowen's Disease: a four-year retrospective review of epidemiology and treatment at a university center. *Dermatol Surg,* 34,(7)**,** 878-83.

Hart, P. H., Gorman, S. & Finlay-Jones, J. J. (2011). Modulation of the immune system by UV radiation: more than just the effects of vitamin D? *Nat Rev Immunol,* 11,(9)**,** 584-96.

Hartevelt, M. M., Bavinck, J. N., Kootte, A. M., Vermeer, B. J. & Vandenbroucke, J. P. (1990). Incidence of skin cancer after renal transplantation in The Netherlands. *Transplantation,* 49,(3)**,** 506-9.

Harvey, I., Frankel, S., Marks, R., Shalom, D. & Nolan-Farrell, M. (1996). Non-melanoma skin cancer and solar keratoses. I. Methods and descriptive results of the South Wales Skin Cancer Study. *Br J Cancer,* 74,(8)**,** 1302-7.

Harwood, C. A., Attard, N. R., O'Donovan, P., Chambers, P., Perrett, C. M., Proby, C. M., McGregor, J. M. & Karran, P. (2008). PTCH mutations in basal cell carcinomas from azathioprine-treated organ transplant recipients. *Br J Cancer,* 99,(8)**,** 1276-84.

Healy, E., Flannagan, N., Ray, A., Todd, C., Jackson, I. J., Matthews, J. N., Birch-Machin, M. A. & Rees, J. L. (2000). Melanocortin-1-receptor gene and sun sensitivity in individuals without red hair. *Lancet,* 355,(9209)**,** 1072-3.

Healy, E., Jordan, S. A., Budd, P. S., Suffolk, R., Rees, J. L. & Jackson, I. J. (2001). Functional variation of MC1R alleles from red-haired individuals. *Hum Mol Genet,* 10,(21)**,** 2397-402.

Heaphy, M. R. & Ackerman, A. B. (2000). The nature of solar keratosis: a critical review in historical perspective. *J Am Acad Dermatol,* 43,(1)**,** 138-50.

Heitzer, E., Lassacher, A., Quehenberger, F., Kerl, H. & Wolf, P. (2007). UV fingerprints predominate in the PTCH mutation spectra of basal cell carcinomas independent of clinical phenotype. *J Invest Dermatol,* 127,(12)**,** 2872-81.

Hobbs, G. A., Der, C. J. & Rossman, K. L. (2016). RAS isoforms and mutations in cancer at a glance. *J Cell Sci,* 129,(7)**,** 1287-92.

Hodis, E., Watson, I. R., Kryukov, G. V., Arold, S. T., Imielinski, M., Theurillat, J. P., Nickerson, E., Auclair, D., Li, L., Place, C., Dicara, D., Ramos, A. H., Lawrence, M. S., Cibulskis, K., Sivachenko, A., Voet, D., Saksena, G., Stransky, N., Onofrio, R. C., Winckler, W., et al. (2012). A landscape of driver mutations in melanoma. *Cell,* 150,(2)**,** 251-63.

Hoffman, W. H., Biade, S., Zilfou, J. T., Chen, J. & Murphy, M. (2002). Transcriptional repression of the anti-apoptotic survivin gene by wild type p53. *J Biol Chem,* 277,(5)**,** 3247-57.

Hollstein, M., Sidransky, D., Vogelstein, B. & Harris, C. C. (1991). p53 mutations in human cancers. *Science,* 253,(5015)**,** 49-53.

Hopfer, O., Zwahlen, D., Fey, M. F. & Aebi, S. (2005). The Notch pathway in ovarian carcinomas and adenomas. *Br J Cancer,* 93,(6)**,** 709-18.

Hostetter, G., Kim, S. Y., Savage, S., Gooden, G. C., Barrett, M., Zhang, J., Alla, L., Watanabe, A., Einspahr, J., Prasad, A., Nickoloff, B. J., Carpten, J., Trent, J., Alberts, D. & Bittner, M. (2010). Random DNA fragmentation allows detection of single-copy, single-exon alterations of copy number by oligonucleotide array CGH in clinical FFPE samples. *Nucleic Acids Res,* 38,(2)**,** 1-12.

Howard, M. & O'Garra, A. (1992). Biological properties of interleukin 10. *Immunol Today,* 13,(6)**,** 198-200.

Hu, B., Castillo, E., Harewood, L., Ostano, P., Reymond, A., Dummer, R., Raffoul, W., Hoetzenecker, W., Hofbauer, G. F. & Dotto, G. P. (2012). Multifocal epithelial tumors and field cancerization from loss of mesenchymal CSL signaling. *Cell,* 149,(6)**,** 1207-20.

Huang, Y. S., Bu, D. F., Li, X. Y., Ma, Z. H., Yang, Y., Lin, Z. M., Lu, F. M., Tu, P. & Li, H. (2013). Unique features of PTCH1 mutation spectrum in Chinese sporadic basal cell carcinoma. *J Eur Acad Dermatol Venereol,* 27,(2)**,** 235-41.

Huber, A. H., Nelson, W. J. & Weis, W. I. (1997). Three-dimensional structure of the armadillo repeat region of beta-catenin. *Cell,* 90,(5)**,** 871-82.

Hunter, J. C., Manandhar, A., Carrasco, M. A., Gurbani, D., Gondi, S. & Westover, K. D. (2015). Biochemical and structural analysis of common cancer-associated KRAS mutations. *Mol Cancer Res,* 13,(9)**,** 1325-35.

Illumina. *DesignStudio* [Online]. Available: www.designstudio.illumina.com. [Accessed 07 April 2015].

Illumina. (2015). *TruSeq® Custom Amplicon v1.5: A new and improved amplicon sequencing solution for interrogating custom regions of interest.* [Online]. Available: http://www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/datasheet_truseq_custom_amplicon.pdf. [Accessed 28 August 2015].

Ishida, H., Kumakiri, M., Ueda, K., Lao, L. M., Yanagihara, M., Asamoto, K., Imamura, Y., Noriki, S. & Fukuda, M. (2001). Comparative histochemical study of Bowen's disease and actinic keratosis: preserved normal basal cells in Bowen's disease. *Eur J Histochem,* 45,(2)**,** 177-90.

Isobe, M., Emanuel, B. S., Givol, D., Oren, M. & Croce, C. M. (1986). Localization of gene for human p53 tumour antigen to band 17p13. *Nature,* 320,(6057)**,** 84-5.

Ito, S. & Wakamatsu, K. (2003). Quantitative analysis of eumelanin and pheomelanin in humans, mice, and other animals: a comparative review. *Pigment Cell Res,* 16,(5)**,** 523-31.

Jacobs, G. H., Rippey, J. J. & Altini, M. (1982). Prediction of aggressive behavior in basal cell carcinoma. *Cancer,* 49,(3)**,** 533-7.

Jaeger, A. B., Gramkow, A., Hjalgrim, H., Melbye, M. & Frisch, M. (1999). Bowen disease and risk of subsequent malignant neoplasms: a population-based cohort study of 1147 patients. *Arch Dermatol,* 135,(7)**,** 790-3.

Jang, T. J. (2008). Prevalence of Foxp3 positive T regulatory cells is increased during progression of cutaneous squamous tumors. *Yonsei Med J,* 49,(6)**,** 942-8.

Jayaraman, S. S., Rayhan, D. J., Hazany, S. & Kolodney, M. S. (2014). Mutational landscape of basal cell carcinomas by whole-exome sequencing. *J Invest Dermatol,* 134,(1)**,** 213-20.

Jeffers, J. R., Parganas, E., Lee, Y., Yang, C., Wang, J., Brennan, J., MacLean, K. H., Han, J., Chittenden, T., Ihle, J. N., McKinnon, P. J., Cleveland, J. L. & Zambetti, G. P. (2003). Puma is an essential mediator of p53-dependent and -independent apoptotic pathways. *Cancer Cell,* 4,(4)**,** 321-8.

Jemal, A., Siegel, R., Xu, J. & Ward, E. (2010). Cancer statistics, 2010. *CA Cancer J Clin,* 60,(5)**,** 277-300.

Jensen, P., Hansen, S., Moller, B., Leivestad, T., Pfeffer, P., Geiran, O., Fauchald, P. & Simonsen, S. (1999). Skin cancer in kidney and heart transplant recipients and different long-term immunosuppressive therapy regimens. *J Am Acad Dermatol,* 40,(2)**,** 177-86.

Jiang, W., Ananthaswamy, H. N., Muller, H. K. & Kripke, M. L. (1999). p53 protects against skin cancer induction by UV-B radiation. *Oncogene,* 18,(29)**,** 4247-53.

Jonason, A. S., Kunala, S., Price, G. J., Restifo, R. J., Spinelli, H. M., Persing, J. A., Leffell, D. J., Tarone, R. E. & Brash, D. E. (1996). Frequent clones of p53-mutated keratinocytes in normal human skin. *Proc Natl Acad Sci U S A,* 93,(24)**,** 14025-9.

Jung, G. W., Metelitsa, A. I., Dover, D. C. & Salopek, T. G. (2010). Trends in incidence of nonmelanoma skin cancers in Alberta, Canada, 1988-2007. *Br J Dermatol,* 163,(1)**,** 146-54.

Kamb, A., Gruis, N. A., Weaver-Feldhaus, J., Liu, Q., Harshman, K., Tavtigian, S. V., Stockert, E., Day, R. S., Johnson, B. E. & Skolnick, M. H. (1994). A cell cycle regulator potentially involved in genesis of many tumor types. *Science,* 264,(5157)**,** 436-40.

Kan, Z., Jaiswal, B. S., Stinson, J., Janakiraman, V., Bhatt, D., Stern, H. M., Yue, P., Haverty, P. M., Bourgon, R., Zheng, J., Moorhead, M., Chaudhuri, S., Tomsho, L. P., Peters, B. A., Pujara, K., Cordes, S., Davis, D. P., Carlton, V. E., Yuan, W., Li, L., et al. (2010). Diverse somatic mutation patterns and pathway alterations in human cancers. *Nature,* 466,(7308)**,** 869-73.

Kandoth, C., McLellan, M. D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J. F., Wyczalkowski, M. A., Leiserson, M. D., Miller, C. A., Welch, J. S., Walter, M. J., Wendl, M. C., Ley, T. J., Wilson, R. K., Raphael, B. J. & Ding, L. (2013). Mutational landscape and significance across 12 major cancer types. *Nature,* 502,(7471)**,** 333-9.

Kanellou, P., Zaravinos, A., Zioga, M., Stratigos, A., Baritaki, S., Soufla, G., Zoras, O. & Spandidos, D. A. (2008). Genomic instability, mutations and expression analysis of the tumour suppressor genes p14(ARF), p15(INK4b), p16(INK4a) and p53 in actinic keratosis. *Cancer Lett,* 264,(1)**,** 145-61.

Kanjilal, S., Strom, S. S., Clayman, G. L., Weber, R. S., El-Naggar, A. K., Kapur, V., Cummings, K. K., Hill, L. A., Spitz, M. R., Kripke, M. L. & Ananthaswamy, H. N. (1995). p53 mutations in nonmelanoma skin cancer of the head and neck: molecular evidence for field cancerization. *Cancer Res,* 55,(16)**,** 3604-9.

Kao, G. F. (1986). Carcinoma arising in Bowen's disease. *Arch Dermatol,* 122,(10)**,** 1124-6.

Karagas, M. R., Stannard, V. A., Mott, L. A., Slattery, M. J., Spencer, S. K. & Weinstock, M. A. (2002). Use of tanning devices and risk of basal cell and squamous cell skin cancers. *J Natl Cancer Inst,* 94,(3)**,** 224-6.

Kasper, M., Jaks, V., Hohl, D. & Toftgard, R. (2012). Basal cell carcinoma - molecular biology and potential new therapies. *J Clin Invest,* 122,(2)**,** 455-63.

Kelly, D. A., Young, A. R., McGregor, J. M., Seed, P. T., Potten, C. S. & Walker, S. L. (2000). Sensitivity to sunburn is associated with susceptibility to ultraviolet radiation-induced suppression of cutaneous cell-mediated immunity. *J Exp Med,* 191,(3)**,** 561-6.

Kemp, C. J., Donehower, L. A., Bradley, A. & Balmain, A. (1993). Reduction of p53 gene dosage does not increase initiation or promotion but enhances malignant progression of chemically induced skin tumors. *Cell,* 74,(5)**,** 813-22.

Kennedy, C., Ter Huurne, J., Berkhout, M., Gruis, N., Bastiaens, M., Bergman, W., Willemze, R. & Bavinck, J. N. (2001). Melanocortin 1 receptor (MC1R) gene variants are associated with an increased risk for cutaneous melanoma which is largely independent of skin type and hair color. *J Invest Dermatol,* 117,(2)**,** 294-300.

Kiel, C., Wohlgemuth, S., Rousseau, F., Schymkowitz, J., Ferkinghoff-Borg, J., Wittinghofer, F. & Serrano, L. (2005). Recognizing and defining true Ras binding domains II: in silico prediction based on homology modelling and energy calculations. *J Mol Biol,* 348,(3)**,** 759-75.

Kim, M. Y., Park, H. J., Baek, S. C., Byun, D. G. & Houh, D. (2002). Mutations of the p53 and PTCH gene in basal cell carcinomas: UV mutation signature and strand bias. *J Dermatol Sci,* 29,(1)**,** 1-9.

Klein, C. A. (2009). Parallel progression of primary tumours and metastases. *Nat Rev Cancer,* 9,(4)**,** 302-12.

Ko, C. J. (2010). Actinic keratosis: facts and controversies. *Clin Dermatol,* 28,(3)**,** 249-53.

Koboldt, D. C., Zhang, Q., Larson, D. E., Shen, D., McLellan, M. D., Lin, L., Miller, C. A., Mardis, E. R., Ding, L. & Wilson, R. K. (2012). VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res,* 22,(3)**,** 568-76.

Koseki, S., Aoki, T., Ansai, S., Hozumi, Y., Mitsuhashi, Y. & Kondo, S. (1999). An immunohistochemical study of E-cadherin expression in human squamous cell carcinoma of the skin: relationship between decreased expression of E-cadherin in the primary lesion and regional lymph node metastasis. *J Dermatol,* 26,(7)**,** 416-22.

Kossard, S. & Rosen, R. (1992). Cutaneous Bowen's disease. An analysis of 1001 cases according to age, sex, and site. *J Am Acad Dermatol,* 27,(3)**,** 406-10.

Kovacs, A., Yonemoto, K., Katsuoka, K., Nishiyama, S. & Harhai, I. (1996). Bowen's disease: statistical study of a 10 year period. *J Dermatol,* 23,(4)**,** 267-74.

Kramata, P., Lu, Y. P., Lou, Y. R., Singh, R. N., Kwon, S. M. & Conney, A. H. (2005). Patches of mutant p53-immunoreactive epidermal cells induced by chronic UVB Irradiation harbor the same p53 mutations as squamous cell carcinomas in the skin of hairless SKH-1 mice. *Cancer Res,* 65,(9)**,** 3577-85.

Kraus, D. H., Carew, J. F. & Harrison, L. B. (1998). Regional lymph node metastasis from cutaneous squamous cell carcinoma. *Arch Otolaryngol Head Neck Surg,* 124,(5)**,** 582-7.

Kress, S., Sutter, C., Strickland, P. T., Mukhtar, H., Schweizer, J. & Schwarz, M. (1992). Carcinogen-specific mutational pattern in the p53 gene in ultraviolet B radiation-induced squamous cell carcinomas of mouse skin. *Cancer Res,* 52,(22)**,** 6400-3.

Kricker, A., Armstrong, B. K., English, D. R. & Heenan, P. J. (1995). Does intermittent sun exposure cause basal cell carcinoma? a case-control study in Western Australia. *Int J Cancer,* 60,(4)**,** 489-94.

Kubo, Y., Urano, Y., Yoshimoto, K., Iwahana, H., Fukuhara, K., Arase, S. & Itakura, M. (1994). p53 gene mutations in human skin cancers and precancerous lesions: comparison with immunohistochemical analysis. *J Invest Dermatol,* 102,(4)**,** 440-4.

Kushida, Y., Miki, H. & Ohmori, M. (1999). Loss of heterozygosity in actinic keratosis, squamous cell carcinoma and sun-exposed normal-appearing skin in Japanese: difference between Japanese and Caucasians. *Cancer Lett,* 140,(1)**,** 169-75.

Küsters-Vandevelde, H. V., Van Leeuwen, A., Verdijk, M. A., De Koning, M. N., Quint, W. G., Melchers, W. J., Ligtenberg, M. J. & Blokx, W. A. (2010). CDKN2A but not TP53 mutations nor HPV presence predict poor outcome in metastatic squamous cell carcinoma of the skin. *Int J Cancer,* 126,(9)**,** 2123-32.

Kwa, R. E., Campana, K. & Moy, R. L. (1992). Biology of cutaneous squamous cell carcinoma. *J Am Acad Dermatol,* 26,(1)**,** 1-26.

Lai, C., August, S., Albibas, A., Behar, R., Cho, S. Y., Polak, M. E., Theaker, J., MacLeod, A. S., French, R. R., Glennie, M. J., Al-Shamkhani, A. & Healy, E. (2016). OX40+ regulatory T cells in cutaneous squamous cell carcinoma suppress effector T-cell responses and associate with metastatic potential. *Clin Cancer Res,* 22,(16)**,** 4236-48.

Lakin, N. D. & Jackson, S. P. (1999). Regulation of p53 in response to DNA damage. *Oncogene,* 18,(53)**,** 7644-55.

Lamar, E., Deblandre, G., Wettstein, D., Gawantka, V., Pollet, N., Niehrs, C. & Kintner, C. (2001). Nrarp is a novel intracellular component of the Notch signaling pathway. *Genes Dev,* 15,(15)**,** 1885-99.

Lamb, P. & Crawford, L. (1986). Characterization of the human p53 gene. *Mol Cell Biol,* 6,(5)**,** 1379-85.

Lane, D. P. & Crawford, L. V. (1979). T antigen is bound to a host protein in SV40-transformed cells. *Nature,* 278,(5701)**,** 261-3.

Lapouge, G., Youssef, K. K., Vokaer, B., Achouri, Y., Michaux, C., Sotiropoulou, P. A. & Blanpain, C. (2011). Identifying the cellular origin of squamous skin tumors. *Proc Natl Acad Sci U S A,* 108,(18)**,** 7431-6.

Latonen, L. & Laiho, M. (2005). Cellular UV damage responses--functions of tumor suppressor p53. *Biochim Biophys Acta,* 1755,(2)**,** 71-89.

Lawrence, M. S., Stojanov, P., Polak, P., Kryukov, G. V., Cibulskis, K., Sivachenko, A., Carter, S. L., Stewart, C., Mermel, C. H., Roberts, S. A., Kiezun, A., Hammerman, P. S., McKenna, A., Drier, Y., Zou, L., Ramos, A. H., Pugh, T. J., Stransky, N., Helman, E., Kim, J., et al. (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature,* 499,(7457)**,** 214-8.

Lee, C. S., Bhaduri, A., Mah, A., Johnson, W. L., Ungewickell, A., Aros, C. J., Nguyen, C. B., Rios, E. J., Siprashvili, Z., Straight, A., Kim, J., Aasi, S. Z. & Khavari, P. A. (2014). Recurrent point mutations in the kinetochore gene KNSTRN in cutaneous squamous cell carcinoma. *Nat Genet,* 46,(10)**,** 1060-2.

Lee, H. J., Kim, J. S., Ha, S. J., Roh, K. Y., Seo, E. J., Park, W. S., Lee, J. Y., Park, K. S. & Kim, J. W. (2000). p53 gene mutations in Bowen's disease in Koreans: clustering in exon 5 and multiple mutations. *Cancer Lett,* 158,(1)**,** 27-33.

Leffell, D. J. (2000). The scientific basis of skin cancer. *J Am Acad Dermatol,* 42,(1)**,** 18-22.

Lefort, K., Mandinova, A., Ostano, P., Kolev, V., Calpini, V., Kolfschoten, I., Devgan, V., Lieb, J., Raffoul, W., Hohl, D., Neel, V., Garlick, J., Chiorino, G. & Dotto, G. P. (2007). Notch1 is a p53 target gene involved in human keratinocyte tumor suppression through negative regulation of ROCK1/2 and MRCKalpha kinases. *Genes Dev,* 21,(5)**,** 562-77.

Lehman, T. A., Modali, R., Boukamp, P., Stanek, J., Bennett, W. P., Welsh, J. A., Metcalf, R. A., Stampfer, M. R., Fusenig, N., Rogan, E. M. & Harris, C. C. (1993). p53 mutations in human immortalized epithelial cell lines. *Carcinogenesis,* 14,(5)**,** 833-9.

Leibovitch, I., Huilgol, S. C., Selva, D., Richards, S. & Paver, R. (2005). Cutaneous squamous carcinoma in situ (Bowen's disease): treatment with Mohs micrographic surgery. *J Am Acad Dermatol,* 52,(6)**,** 997-1002.

Leis, P. F., Stevens, K. R., Baer, S. C., Kadmon, D., Goldberg, L. H. & Wang, X. J. (1998). A c-rasHa mutation in the metastasis of a human papillomavirus (HPV)-18 positive penile squamous cell carcinoma suggests a cooperative effect between HPV-18 and c-rasHa activation in malignant progression. *Cancer,* 83,(1)**,** 122-9.

Leonardi-Bee, J., Ellison, T. & Bath-Hextall, F. (2012). Smoking and the risk of nonmelanoma skin cancer: Systematic review and meta-analysis. *Arch Dermatol,* 148,(8)**,** 939-46.

Levitan, D., Lee, J., Song, L., Manning, R., Wong, G., Parker, E. & Zhang, L. (2001). PS1 N- and C-terminal fragments form a complex that functions in APP processing and Notch signaling. *Proc Natl Acad Sci U S A,* 98,(21)**,** 12186-90.

Li, Y. Y., Hanna, G. J., Laga, A. C., Haddad, R. I., Lorch, J. H. & Hammerman, P. S. (2015). Genomic analysis of metastatic cutaneous squamous cell carcinoma. *Clin Cancer Res,* 21,(6)**,** 1447-56.

Lin, J. Y. & Fisher, D. E. (2007). Melanocyte biology and skin pigmentation. *Nature,* 445,(7130)**,** 843-50.

Lin, W., Qureshi, A. A., Kraft, P., Nan, H., Guo, Q., Hu, F. B., Jensen, M. K. & Han, J. (2011). ASIP genetic variants and the number of non-melanoma skin cancers. *Cancer Causes Control,* 22,(3)**,** 495-501.

Liu, C., Li, Y., Semenov, M., Han, C., Baeg, G. H., Tan, Y., Zhang, Z., Lin, X. & He, X. (2002). Control of beta-catenin phosphorylation/degradation by a dual-kinase mechanism. *Cell,* 108,(6)**,** 837-47.

Lober, B. A. & Lober, C. W. (2000). Actinic keratosis is squamous cell carcinoma. *South Med J,* 93,(7)**,** 650-5.

Loeb, K. R., Asgari, M. M., Hawes, S. E., Feng, Q., Stern, J. E., Jiang, M., Argenyi, Z. B., De Villiers, E. M. & Kiviat, N. B. (2012). Analysis of Tp53 codon 72 polymorphisms, Tp53 mutations, and HPV infection in cutaneous squamous cell carcinomas. *PLOS One,* 7,(4)**,** 1-11.

Loeb, L. A. (1991). Mutator phenotype may be required for multistage carcinogenesis. *Cancer Res,* 51,(12)**,** 3075-9.

Lohmann, C. M. & Solomon, A. R. (2001). Clinicopathologic variants of cutaneous squamous cell carcinoma. *Adv Anat Pathol,* 8,(1)**,** 27-36.

Lomas, A., Leonardi-Bee, J. & Bath-Hextall, F. (2012). A systematic review of worldwide incidence of nonmelanoma skin cancer. *Br J Dermatol,* 166,(5)**,** 1069-80.

Lu, Y. P., Lou, Y. R., Liao, J., Xie, J. G., Peng, Q. Y., Yang, C. S. & Conney, A. H. (2005). Administration of green tea or caffeine enhances the disappearance of UVB-induced

patches of mutant p53 positive epidermal cells in SKH-1 mice. *Carcinogenesis,* 26,(8)**,** 1465-72.

Lucas, R., McMichael, T., Smith, W. & Armstrong, B. (2006). *Solar Ultraviolet Radiation: Global burden of disease from solar ultraviolet radiation.* Geneva: WHO.

Ludyga, N., Grunwald, B., Azimzadeh, O., Englert, S., Hofler, H., Tapio, S. & Aubele, M. (2012). Nucleic acids from long-term preserved FFPE tissues are suitable for downstream analyses. *Virchows Arch,* 460,(2)**,** 131-40.

Mabula, J. B., Chalya, P. L., McHembe, M. D., Jaka, H., Giiti, G., Rambau, P., Masalu, N., Kamugisha, E., Robert, S. & Gilyoma, J. M. (2012). Skin cancers among Albinos at a University teaching hospital in Northwestern Tanzania: a retrospective review of 64 cases. *BMC Dermatol,* 12**,** 5-12.

Madan, V., Lear, J. T. & Szeimies, R. M. (2010). Non-melanoma skin cancer. *Lancet,* 375,(9715)**,** 673-85.

Malumbres, M. & Pellicer, A. (1998). RAS pathways to cell cycle control and cell transformation. *Front Biosci,* 3,(6)**,** 887-912.

Mao, L., Merlo, A., Bedi, G., Shapiro, G. I., Edwards, C. D., Rollins, B. J. & Sidransky, D. (1995). A novel p16INK4A transcript. *Cancer Res,* 55,(14)**,** 2995-7.

Marks, R., Foley, P., Goodman, G., Hage, B. H. & Selwood, T. S. (1986). Spontaneous remission of solar keratoses: the case for conservative management. *Br J Dermatol,* 115,(6)**,** 649-55.

Marks, R., Rennie, G. & Selwood, T. S. (1988). Malignant transformation of solar keratoses to squamous cell carcinoma. *Lancet,* 1,(8589)**,** 795-7.

Marshall, C. J., Hall, A. & Weiss, R. A. (1982). A transforming gene present in human sarcoma cell lines. *Nature,* 299,(5879)**,** 171-3.

Martincorena, I., Roshan, A., Gerstung, M., Ellis, P., Van Loo, P., McLaren, S., Wedge, D. C., Fullam, A., Alexandrov, L. B., Tubio, J. M., Stebbings, L., Menzies, A., Widaa, S., Stratton, M. R., Jones, P. H. & Campbell, P. J. (2015). High burden and pervasive positive selection of somatic mutations in normal human skin. *Science,* 348,(6237)**,** 880-6.

Martinez-Zapien, D., Ruiz, F. X., Poirson, J., Mitschler, A., Ramirez, J., Forster, A., Cousido-Siah, A., Masson, M., Vande-Pol, S., Podjarny, A., Trave, G. & Zanier, K. (2016). Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53. *Nature,* 529,(7587)**,** 541-5.

Martinez, J. C., Otley, C. C., Stasko, T., Euvrard, S., Brown, C., Schanbacher, C. F. & Weaver, A. L. (2003). Defining the clinical course of metastatic skin cancer in organ transplant recipients: a multicenter collaborative study. *Arch Dermatol,* 139,(3)**,** 301-6.

Matsuda, T., Bebenek, K., Masutani, C., Hanaoka, F. & Kunkel, T. A. (2000). Low fidelity DNA synthesis by human DNA polymerase-η. *Nature,* 404,(6781)**,** 1011-3.

Mauerer, A., Herschberger, E., Dietmaier, W., Landthaler, M. & Hafner, C. (2011). Low incidence of EGFR and HRAS mutations in cutaneous squamous cell carcinomas of a German cohort. *Exp Dermatol,* 20,(10)**,** 848-50.

May, P. & May, E. (1999). Twenty years of p53 research: structural and functional aspects of the p53 protein. *Oncogene,* 18,(53)**,** 7621-36.

McCusker, M., Basset-Seguin, N., Dummer, R., Lewis, K., Schadendorf, D., Sekulic, A., Hou, J., Wang, L., Yue, H. & Hauschild, A. (2014). Metastatic basal cell carcinoma: Prognosis dependent on anatomic site and spread of disease. *Eur J Cancer,* 50,(4)**,** 774-83.

McGregor, J. M., Berkhout, R. J., Rozycka, M., Ter Schegget, J., Bouwes-Bavinck, J. N., Brooks, L. & Crook, T. (1997). p53 mutations implicate sunlight in post-transplant skin cancer irrespective of human papillomavirus status. *Oncogene,* 15,(14)**,** 1737-40.

Meldrum, C., Doyle, M. A. & Tothill, R. W. (2011). Next-generation sequencing for cancer diagnostics: a practical perspective. *Clin Biochem Rev,* 32,(4)**,** 177-95.

Melnikova, V. O., Pacifico, A., Chimenti, S., Peris, K. & Ananthaswamy, H. N. (2005). Fate of UVB-induced p53 mutations in SKH-hr1 mouse skin after discontinuation of irradiation: relationship to skin cancer development. *Oncogene,* 24,(47)**,** 7055-63.

Merlo, L. M., Pepper, J. W., Reid, B. J. & Maley, C. C. (2006). Cancer as an evolutionary and ecological process. *Nat Rev Cancer,* 6,(12)**,** 924-35.

Micali, G., Lacarrubba, F., Bongu, A. & West, D. (2001). The skin barrier. *In:* Woodly, D. T. & Freinkel, R. K. (eds.) IN: *The biology of the skin.* New York: The Parthoenon Publishing Group Inc.

Michel, A., Kopp-Schneider, A., Zentgraf, H., Gruber, A. D. & De Villiers, E. M. (2006). E6/E7 expression of human papillomavirus type 20 (HPV-20) and HPV-27 influences proliferation and differentiation of the skin in UV-irradiated SKH-hr1 transgenic mice. *J Virol,* 80,(22)**,** 11153-64.

Miller, J. H. (1985). Mutagenic specificity of ultraviolet light. *J Mol Biol,* 182,(1)**,** 45-65.

Milon, A., Bulliard, J. L., Vuilleumier, L., Danuser, B. & Vernez, D. (2014). Estimating the contribution of occupational solar ultraviolet exposure to skin cancer. *Br J Dermatol,* 170,(1)**,** 157-64.

Mittelbronn, M. A., Mullins, D. L., Ramos-Caro, F. A. & Flowers, F. P. (1998). Frequency of pre-existing actinic keratosis in cutaneous squamous cell carcinoma. *Int J Dermatol,* 37,(9)**,** 677-81.

Moles, J. P., Moyret, C., Guillot, B., Jeanteur, P., Guilhou, J. J., Theillet, C. & Basset-Seguin, N. (1993). p53 gene mutations in human epithelial skin cancers. *Oncogene,* 8,(3)**,** 583-8.

Moll, I., Roessler, M., Brandner, J. M., Eispert, A. C., Houdek, P. & Moll, R. (2005). Human Merkel cells--aspects of cell biology, distribution and functions. *Eur J Cell Biol,* 84,(2)**,** 259-71.

Mori, H., Colman, S. M., Xiao, Z., Ford, A. M., Healy, L. E., Donaldson, C., Hows, J. M., Navarrete, C. & Greaves, M. (2002). Chromosome translocations and covert leukemic clones are generated during normal fetal development. *Proc Natl Acad Sci U S A,* 99,(12)**,** 8242-7.

Morin, R. D., Mendez-Lago, M., Mungall, A. J., Goya, R., Mungall, K. L., Corbett, R. D., Johnson, N. A., Severson, T. M., Chiu, R., Field, M., Jackman, S., Krzywinski, M., Scott, D. W., Trinh, D. L., Tamura-Wells, J., Li, S., Firme, M. R., Rogic, S., Griffith, M., Chan, S., et al. (2011). Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature,* 476,(7360)**,** 298-303.

Morris, S., Cox, B. & Bosanquet, N. (2009). Cost of skin cancer in England. *Eur J Health Econ,* 10,(3)**,** 267-73.

Mortier, L., Marchetti, P., Delaporte, E., De Lassalle, E. M., Thomas, P., Piette, F., Formstecher, P., Polakowska, R. & Danze, P. M. (2002). Progression of actinic keratosis to squamous cell carcinoma of the skin correlates with deletion of the 9p21 region encoding the p16(INK4a) tumor suppressor. *Cancer Lett,* 176,(2)**,** 205-14.

Morton, C. A., Birnie, A. J. & Eedy, D. J. (2014). British Association of Dermatologists' guidelines for the management of squamous cell carcinoma in situ (Bowen's disease) 2014. *Br J Dermatol,* 170,(2)**,** 245-60.

Moy, R. L. (2000). Clinical presentation of actinic keratoses and squamous cell carcinoma. *J Am Acad Dermatol,* 42,(Supp 1)**,** S8-10.

Murao, K., Yoshioka, R. & Kubo, Y. (2014). Human papillomavirus infection in Bowen disease: negative p53 expression, not p16(INK4a) overexpression, is correlated with human papillomavirus-associated Bowen disease. *J Dermatol,* 41,(10)**,** 878-84.

Murata, Y., Kumano, K. & Sashikata, T. (1996). Partial spontaneous regression of Bowen's disease. *Arch Dermatol,* 132,(4)**,** 429-32.

Nan, H., Kraft, P., Hunter, D. J. & Han, J. (2009). Genetic variants in pigmentation genes, pigmentary phenotypes, and risk of skin cancer in Caucasians. *Int J Cancer,* 125,(4)**,** 909-17.

Nassar, D., Latil, M., Boeckx, B., Lambrechts, D. & Blanpain, C. (2015). Genomic landscape of carcinogen-induced and genetically induced mouse skin squamous cell carcinoma. *Nat Med,* 21,(8)**,** 946-54.

NCBI. *Gene.* [Online]. Available: www.ncbi.nlm.nih.gov/gene. [Accessed 03 September 2016].

NCBI. *Protein.* [Online]. Available: www.ncbi.nlm.nih.gov/protein/. [Accessed 03 May 2016].

NCBI National Center for Biotechnology Information. *BLAST.* [Online]. Available: www.blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome. [Accessed 20 August 2016].

NCG5.0 Network of Cancer Gene. [Online]. Available: www.ncg.kcl.ac.uk. [Accessed 17 February 2015].

Nelson, M. A., Einspahr, J. G., Alberts, D. S., Balfour, C. A., Wymer, J. A., Welch, K. L., Salasche, S. J., Bangert, J. L., Grogan, T. M. & Bozzo, P. O. (1994). Analysis of the p53 gene in human precancerous actinic keratosis lesions and squamous cell cancers. *Cancer Lett,* 85,(1)**,** 23-9.

Nestle, F. O., Di Meglio, P., Qin, J. Z. & Nickoloff, B. J. (2009). Skin immune sentinels in health and disease. *Nat Rev Immunol,* 9,(10)**,** 679-91.

Neto, P. D., Alchorne, M., Michalany, N., Abreu, M. & Borra, R. (2013). Reduced p53 staining in actinic keratosis is associated with squamous cell carcinoma: A preliminary study. *Indian J Dermatol,* 58,(4)**,** 325-9.

Nguyen, B. C., Lefort, K., Mandinova, A., Antonini, D., Devgan, V., Della-Gatta, G., Koster, M. I., Zhang, Z., Wang, J., Di Vignano, A. T., Kitajewski, J., Chiorino, G., Roop, D. R., Missero, C. & Dotto, G. P. (2006). Cross-regulation between Notch and p63 in keratinocyte commitment to differentiation. *Genes Dev,* 20,(8)**,** 1028-42.

Nicolas, M., Wolfer, A., Raj, K., Kummer, J. A., Mill, P., Van Noort, M., Hui, C. C., Clevers, H., Dotto, G. P. & Radtke, F. (2003). Notch1 functions as a tumor suppressor in mouse skin. *Nat Genet,* 33,(3)**,** 416-21.

Nihei, N., Hiruma, M., Ikeda, S. & Ogawa, H. (2004). A case of Bowen's disease showing a clinical tendency toward spontaneous regression. *J Dermatol,* 31,(7)**,** 569-72.

Nikolaev, S. I., Rimoldi, D., Iseli, C., Valsesia, A., Robyr, D., Gehrig, C., Harshman, K., Guipponi, M., Bukach, O., Zoete, V., Michielin, O., Muehlethaler, K., Speiser, D., Beckmann, J. S., Xenarios, I., Halazonetis, T. D., Jongeneel, C. V., Stevenson, B. J. & Antonarakis, S. E. (2012). Exome sequencing identifies recurrent somatic MAP2K1 and MAP2K2 mutations in melanoma. *Nat Genet,* 44,(2)**,** 133-9.

Nindl, I., Gottschling, M., Krawtchenko, N., Lehmann, M. D., Rowert-Huber, J., Eberle, J., Stockfleth, E. & Forschner, T. (2007). Low prevalence of p53, p16(INK4a) and Ha-ras tumour-specific mutations in low-graded actinic keratosis. *Br J Dermatol,* 156,(Suppl 3)**,** S34-9.

Noonan, F. P. & De Fabo, E. C. (1992). Immunosuppression by ultraviolet B radiation: initiation by urocanic acid. *Immunol Today,* 13,(7)**,** 250-4.

Norval, M., Gilmour, J. W. & Simpson, T. J. (1990). The effect of histamine receptor antagonists on immunosuppression induced by the cis-isomer of urocanic acid. *Photodermatol Photoimmunol Photomed,* 7,(6)**,** 243-8.

Nowell, P. C. (1976). The clonal evolution of tumor cell populations. *Science,* 194,(4260)**,** 23-8.

O'Hayre, M., Vazquez-Prado, J., Kufareva, I., Stawiski, E. W., Handel, T. M., Seshagiri, S. & Gutkind, J. S. (2013). The emerging mutational landscape of G proteins and G-protein-coupled receptors in cancer. *Nat Rev Cancer,* 13,(6)**,** 412-24.

Oberholzer, P. A., Kee, D., Dziunycz, P., Sucker, A., Kamsukom, N., Jones, R., Roden, C., Chalk, C. J., Ardlie, K., Palescandolo, E., Piris, A., MacConaill, L. E., Robert, C., Hofbauer, G. F., McArthur, G. A., Schadendorf, D. & Garraway, L. A. (2012). RAS mutations are associated with the development of cutaneous squamous cell tumors in patients treated with RAF inhibitors. *J Clin Oncol,* 30,(3)**,** 316-21.

Ozawa, M., Baribault, H. & Kemler, R. (1989). The cytoplasmic domain of the cell adhesion molecule uvomorulin associates with three independent proteins structurally related in different species. *EMBO J,* 8,(6)**,** 1711-7.

Pacifico, A., Goldberg, L. H., Peris, K., Chimenti, S., Leone, G. & Ananthaswamy, H. N. (2008). Loss of CDKN2A and p14ARF expression occurs frequently in human nonmelanoma skin cancers. *Br J Dermatol,* 158,(2)**,** 291-7.

Palomero, T., Lim, W. K., Odom, D. T., Sulis, M. L., Real, P. J., Margolin, A., Barnes, K. C., O'Neil, J., Neuberg, D., Weng, A. P., Aster, J. C., Sigaux, F., Soulier, J., Look, A. T., Young, R. A., Califano, A. & Ferrando, A. A. (2006). NOTCH1 directly regulates c-MYC and activates a feed-forward-loop transcriptional network promoting leukemic cell growth. *Proc Natl Acad Sci U S A,* 103,(48)**,** 18261-6.

Panelos, J., Tarantini, F., Paglierani, M., Di Serio, C., Maio, V., Pellerito, S., Pimpinelli, N., Santucci, M. & Massi, D. (2008). Photoexposition discriminates Notch 1 expression in human cutaneous squamous cell carcinoma. *Mod Pathol,* 21,(3)**,** 316-25.

Papadavid, E., Pignatelli, M., Zakynthinos, S., Krausz, T. & Chu, A. C. (2002). Abnormal immunoreactivity of the E-cadherin/catenin (alpha-, beta-, and gamma-) complex in premalignant and malignant non-melanocytic skin tumours. *J Pathol,* 196,(2)**,** 154-62.

Park, W. S., Lee, H. K., Lee, J. Y., Yoo, N. J., Kim, C. S. & Kim, S. H. (1996). p53 mutations in solar keratoses. *Hum Pathol,* 27,(11)**,** 1180-4.

Parry, M., Rose-Zerilli, M. J., Ljungstrom, V., Gibson, J., Wang, J., Walewska, R., Parker, H., Parker, A., Davis, Z., Gardiner, A., McIver-Brown, N., Kalpadakis, C., Xochelli, A., Anagnostopoulos, A., Fazi, C., Gonzalez de Castro, D., Dearden, C., Pratt, G., Rosenquist, R., Ashton-Key, M., et al. (2015). Genetics and prognostication in Splenic Marginal Zone Lymphoma: Revelations from deep sequencing. *Clin Cancer Res,* 21,(18)**,** 4174-83.

Pasparakis, M., Haase, I. & Nestle, F. O. (2014). Mechanisms regulating skin immunity and inflammation. *Nat Rev Immunol,* 14,(5)**,** 289-301.

Pastore, S., Mascia, F., Mariani, V. & Girolomoni, G. (2008). The epidermal growth factor receptor system in skin repair and inflammation. *J Invest Dermatol,* 128,(6)**,** 1365-74.

Patrick, M. H. (1977). Studies on thymine-derived UV photoproducts in DNA--I. Formation and biological role of pyrimidine adducts in DNA. *Photochem Photobiol,* 25,(4)**,** 357-72.

Pediatric Cancer Genome Project: Explore. (2014). *Protein Painter.* [Online]. Available: www.explore.pediatriccancergenomeproject.org/proteinPainter. [Accessed 23 September 2016].

Pepper, J. W., Findlay, C. S., Kassen, R., Spencer, S. L. & Maley, C. C. (2009). Cancer research meets evolutionary biology. *Evol Appl,* 2,(1)**,** 62-70.

Person, J. R. (2003). An actinic keratosis is neither malignant nor premalignant: it is an initiated tumor. *J Am Acad Dermatol,* 48,(4)**,** 637-8.

Pickering, C. R., Zhou, J. H., Lee, J. J., Drummond, J. A., Peng, S. A., Saade, R. E., Tsai, K. Y., Curry, J. L., Tetzlaff, M. T., Lai, S. Y., Yu, J., Muzny, D. M., Doddapaneni, H., Shinbrot, E., Covington, K. R., Zhang, J., Seth, S., Caulin, C., Clayman, G. L., El-Naggar, A. K., et al. (2014). Mutational landscape of aggressive cutaneous squamous cell carcinoma. *Clin Cancer Res,* 20,(24)**,** 6582-92.

Pierceall, W. E., Kripke, M. L. & Ananthaswamy, H. N. (1992). N-ras mutation in ultraviolet radiation-induced murine skin cancers. *Cancer Res,* 52,(14)**,** 3946-51.

Pierceall, W. E., Mukhopadhyay, T., Goldberg, L. H. & Ananthaswamy, H. N. (1991). Mutations in the p53 tumor suppressor gene in human cutaneous squamous cell carcinomas. *Mol Carcinog,* 4,(6)**,** 445-9.

Ping, X. L., Ratner, D., Zhang, H., Wu, X. L., Zhang, M. J., Chen, F. F., Silvers, D. N., Peacocke, M. & Tsou, H. C. (2001). PTCH mutations in squamous cell carcinoma of the skin. *J Invest Dermatol,* 116,(4)**,** 614-6.

Pleasance, E. D., Cheetham, R. K., Stephens, P. J., McBride, D. J., Humphray, S. J., Greenman, C. D., Varela, I., Lin, M. L., Ordonez, G. R., Bignell, G. R., Ye, K., Alipaz, J., Bauer, M. J., Beare, D., Butler, A., Carter, R. J., Chen, L., Cox, A. J., Edkins, S., Kokko-Gonzales, P. I., et al. (2010). A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature,* 463,(7278)**,** 191-6.

Ponten, F., Berg, C., Ahmadian, A., Ren, Z. P., Nister, M., Lundeberg, J., Uhlen, M. & Ponten, J. (1997). Molecular pathology in basal cell cancer with p53 as a genetic marker. *Oncogene,* 15,(9)**,** 1059-67.

Ponten, F., Berne, B., Ren, Z. P., Nister, M. & Ponten, J. (1995). Ultraviolet light induces expression of p53 and p21 in human skin: effect of sunscreen and constitutive p21 expression in skin appendages. *J Invest Dermatol,* 105,(3)**,** 402-6.

Prickett, T. D., Wei, X., Cardenas-Navia, I., Teer, J. K., Lin, J. C., Walia, V., Gartner, J., Jiang, J., Cherukuri, P. F., Molinolo, A., Davies, M. A., Gershenwald, J. E., Stemke-Hale, K., Rosenberg, S. A., Margulies, E. H. & Samuels, Y. (2011). Exon capture analysis of G protein-coupled receptors identifies activating mutations in GRM3 in melanoma. *Nat Genet,* 43,(11)**,** 1119-26.

Prior, I. A., Lewis, P. D. & Mattos, C. (2012). A comprehensive survey of Ras mutations in cancer. *Cancer Res,* 72,(10)**,** 2457-67.

Quaedvlieg, P. J., Tirsi, E., Thissen, M. R. & Krekels, G. A. (2006). Actinic keratosis: how to differentiate the good from the bad ones? *Eur J Dermatol,* 16,(4)**,** 335-9.

Quelle, D. E., Cheng, M., Ashmun, R. A. & Sherr, C. J. (1997). Cancer-associated mutations at the INK4a locus cancel cell cycle arrest by p16INK4a but not by the alternative reading frame protein p19ARF. *Proc Natl Acad Sci U S A,* 94,(2)**,** 669-73.

Quinlan, A. R. & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics,* 26,(6)**,** 841-2.

Rady, P., Scinicariello, F., Wagner, R. F. & Tyring, S. K. (1992). p53 mutations in basal cell carcinomas. *Cancer Res,* 52,(13)**,** 3804-6.

Ranganathan, P., Weaver, K. L. & Capobianco, A. J. (2011). Notch signalling in solid tumours: a little bit of everything but not all the time. *Nat Rev Cancer,* 11,(5)**,** 338-51.

Rastogi, R. P., Richa, Kumar, A., Tyagi, M. B. & Sinha, R. P. (2010). Molecular mechanisms of ultraviolet radiation-induced DNA damage and repair. *J Nucleic Acids,* 2010**,** 1-32.

Ratushny, V., Gober, M. D., Hick, R., Ridky, T. W. & Seykora, J. T. (2012). From keratinocyte to cancer: the pathogenesis and modeling of cutaneous squamous cell carcinoma. *J Clin Invest,* 122,(2)**,** 464-72.

Rebel, H., Kram, N., Westerman, A., Banus, S., Van Kranen, H. J. & De Gruijl, F. R. (2005). Relationship between UV-induced mutant p53 patches and skin tumours, analysed by mutation spectra and by induction kinetics in various DNA-repair-deficient mice. *Carcinogenesis,* 26,(12)**,** 2123-30.

Rebel, H., Mosnier, L. O., Berg, R. J., De Vries, A. W., Van Steeg, H., Van Kranen, H. J. & De Gruijl, F. R. (2001). Early p53-positive foci as indicators of tumor risk in ultraviolet-exposed hairless mice: kinetics of induction, effects of DNA repair deficiency, and p53 heterozygosity. *Cancer Res,* 61,(3)**,** 977-83.

Rees, J. L. (2004). The genetics of sun sensitivity in humans. *Am J Hum Genet,* 75,(5)**,** 739-51.

Rees, J. L., Birch-Machin, M., Flanagan, N., Healy, E., Phillips, S. & Todd, C. (1999). Genetic studies of the human melanocortin-1 receptor. *Ann N Y Acad Sci,* 885**,** 134-42.

Rehman, I., Quinn, A. G., Healy, E. & Rees, J. L. (1994). High frequency of loss of heterozygosity in actinic keratoses, a usually benign disease. *Lancet,* 344,(8925)**,** 788-9.

Rehman, I., Takata, M., Wu, Y. Y. & Rees, J. L. (1996). Genetic change in actinic keratoses. *Oncogene,* 12,(12)**,** 2483-90.

Reid, B. J., Li, X., Galipeau, P. C. & Vaughan, T. L. (2010). Barrett's oesophagus and oesophageal adenocarcinoma: time for a new synthesis. *Nat Rev Cancer,* 10,(2)**,** 87-101.

Reifenberger, J., Wolter, M., Knobbe, C. B., Kohler, B., Schonicke, A., Scharwachter, C., Kumar, K., Blaschke, B., Ruzicka, T. & Reifenberger, G. (2005). Somatic mutations in the

PTCH, SMOH, SUFUH and TP53 genes in sporadic basal cell carcinomas. *Br J Dermatol,* 152,(1)**,** 43-51.

Reifenberger, J., Wolter, M., Weber, R. G., Megahed, M., Ruzicka, T., Lichter, P. & Reifenberger, G. (1998). Missense mutations in SMOH in sporadic basal cell carcinomas of the skin and primitive neuroectodermal tumors of the central nervous system. *Cancer Res,* 58,(9)**,** 1798-803.

Reizner, G. T., Chuang, T. Y., Elpern, D. J., Stone, J. L. & Farmer, E. R. (1994). Bowen's disease (squamous cell carcinoma in situ) in Kauai, Hawaii. A population-based incidence report. *J Am Acad Dermatol,* 31,(4)**,** 596-600.

Remenyik, E., Wikonkal, N. M., Zhang, W., Paliwal, V. & Brash, D. E. (2003). Antigen-specific immunity does not mediate acute regression of UVB-induced p53-mutant clones. *Oncogene,* 22,(41)**,** 6369-76.

Ren, Z. P., Ahmadian, A., Ponten, F., Nister, M., Berg, C., Lundeberg, J., Uhlen, M. & Ponten, J. (1997). Benign clonal keratinocyte patches with p53 mutations show no genetic link to synchronous squamous cell precancer or cancer in human skin. *Am J Pathol,* 150,(5)**,** 1791-803.

Ren, Z. P., Ponten, F., Nister, M. & Ponten, J. (1996). Two distinct p53 immunohistochemical patterns in human squamous-cell skin cancer, precursors and normal epidermis. *Int J Cancer,* 69,(3)**,** 174-9.

Rigel, D. S. & Stein-Gold, L. F. (2013). The importance of early diagnosis and treatment of actinic keratosis. *J Am Acad Dermatol,* 68,(Suppl 1)**,** S20-7.

Robaina, M. C., Faccion, R. S., Arruda, V. O., De Rezende, L. M., Vasconcelos, G. M., Apa, A. G., Bacchi, C. E. & Klumb, C. E. (2015). Quantitative analysis of CDKN2A methylation, mRNA, and p16(INK4a) protein expression in children and adolescents with Burkitt lymphoma: biological and clinical implications. *Leuk Res,* 39,(2)**,** 248-56.

Robinson, S., Dixon, S., August, S., Diffey, B., Wakamatsu, K., Ito, S., Friedmann, P. S. & Healy, E. (2010). Protection against UVR involves MC1R-mediated non-pigmentary and pigmentary mechanisms in vivo. *J Invest Dermatol,* 130,(7)**,** 1904-13.

Robinson, S. & Healy, E. (2002). Human melanocortin 1 receptor (MC1R) gene variants alter melanoma cell growth and adhesion to extracellular matrix. *Oncogene,* 21,(52)**,** 8037-46.

Rodenhuis, S. (1992). Ras and human tumors. *Semin Cancer Biol,* 3,(4)**,** 241-7.

Rodriguez-Viciana, P., Sabatier, C. & McCormick, F. (2004). Signaling specificity by Ras family GTPases is determined by the full spectrum of effectors they regulate. *Mol Cell Biol,* 24,(11)**,** 4943-54.

Rogers, H. W., Weinstock, M. A., Feldman, S. R. & Coldiron, B. M. (2015). Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the U.S. population, 2012. *JAMA Dermatol,* 151,(10)**,** 1081-6.

Rogers, H. W., Weinstock, M. A., Harris, A. R., Hinckley, M. R., Feldman, S. R., Fleischer, A. B. & Coldiron, B. M. (2010). Incidence estimate of nonmelanoma skin cancer in the United States, 2006. *Arch Dermatol,* 146,(3)**,** 283-7.

Ronchini, C. & Capobianco, A. J. (2001). Induction of cyclin D1 transcription and CDK2 activity by Notch(ic): implication for cell cycle disruption in transformation by Notch(ic). *Mol Cell Biol,* 21,(17)**,** 5925-34.

Rose, R. F., Boon, A., Forman, D., Merchant, W., Bishop, R. & Newton-Bishop, J. A. (2013). An exploration of reported mortality from cutaneous squamous cell carcinoma using death certification and cancer registry data. *Br J Dermatol,* 169,(3)**,** 682-6.

Roshan, A. & Jones, P. H. (2012). Chronic low dose UV exposure and p53 mutation: tilting the odds in early epidermal preneoplasia? *Int J Radiat Biol,* 88,(10)**,** 682-7.

Rosso, S., Zanetti, R., Martinez, C., Tormo, M. J., Schraub, S., Sancho-Garnier, H., Franceschi, S., Gafa, L., Perea, E., Navarro, C., Laurent, R., Schrameck, C., Talamini, R., Tumino, R. & Wechsler, J. (1996). The multicentre south European study 'Helios'. II: Different sun exposure patterns in the aetiology of basal cell and squamous cell carcinomas of the skin. *Br J Cancer,* 73,(11)**,** 1447-54.

Rowe, D. E., Carroll, R. J. & Day, C. L. (1992). Prognostic factors for local recurrence, metastasis, and survival rates in squamous cell carcinoma of the skin, ear, and lip. Implications for treatment modality selection. *J Am Acad Dermatol,* 26,(6)**,** 976-90.

Sage, E. (1993). Distribution and repair of photolesions in DNA: genetic consequences and the role of sequence context. *Photochem Photobiol,* 57,(1)**,** 163-74.

Sah, S., Chen, L., Houghton, J., Kemppainen, J., Marko, A. C., Zeigler, R. & Latham, G. J. (2013). Functional DNA quantification guides accurate next-generation sequencing mutation detection in formalin-fixed, paraffin-embedded tumor biopsies. *Genome Med,* 5,(8)**,** 77-89.

Sakaguchi, S. (2005). Naturally arising Foxp3-expressing CD25+CD4+ regulatory T cells in immunological tolerance to self and non-self. *Nat Immunol,* 6,(4)**,** 345-52.

Sakamoto, K., Fujii, T., Kawachi, H., Miki, Y., Omura, K., Morita, K., Kayamori, K., Katsube, K. & Yamaguchi, A. (2012). Reduction of NOTCH1 expression pertains to maturation abnormalities of keratinocytes in squamous neoplasms. *Lab Invest,* 92,(5)**,** 688-702.

Sakiz, D., Turkmenoglu, T. T. & Kabukcuoglu, F. (2009). The expression of p63 and p53 in keratoacanthoma and intraepidermal and invasive neoplasms of the skin. *Pathol Res Pract,* 205,(9)**,** 589-94.

Salasche, S. J. (2000). Epidemiology of actinic keratoses and squamous cell carcinoma. *J Am Acad Dermatol,* 42,(1)**,** 4-7.

Salo-Mullen, E. E., O'Reilly, E. M., Kelsen, D. P., Ashraf, A. M., Lowery, M. A., Yu, K. H., Reidy, D. L., Epstein, A. S., Lincoln, A., Saldia, A., Jacobs, L. M., Rau-Murthy, R., Zhang, L., Kurtz, R. C., Saltz, L., Offit, K., Robson, M. E. & Stadler, Z. K. (2015). Identification of

germline genetic mutations in patients with pancreatic cancer. *Cancer,* 121,(24)**,** 4382-8.

Samarasinghe, V., Madan, V. & Lear, J. T. (2011). Management of high-risk squamous cell carcinoma of the skin. *Expert Rev Anticancer Ther,* 11,(5)**,** 763-9.

Saridaki, Z., Liloglou, T., Zafiropoulos, A., Koumantaki, E., Zoras, O. & Spandidos, D. A. (2003). Mutational analysis of CDKN2A genes in patients with squamous cell carcinoma of the skin. *Br J Dermatol,* 148,(4)**,** 638-48.

Schmitt, J., Seidler, A., Diepgen, T. L. & Bauer, A. (2011). Occupational ultraviolet light exposure increases the risk for the development of cutaneous squamous cell carcinoma: a systematic review and meta-analysis. *Br J Dermatol,* 164,(2)**,** 291-307.

Schmitt, J. V. & Miot, H. A. (2012). Actinic keratosis: a clinical and epidemiological revision. *An Bras Dermatol,* 87,(3)**,** 425-34.

Schmults, C. D., Karia, P. S., Carter, J. B., Han, J. & Qureshi, A. A. (2013). Factors predictive of recurrence and death from cutaneous squamous cell carcinoma: a 10-year, single-institution cohort study. *JAMA Dermatol,* 149,(5)**,** 541-7.

Scholzen, T. & Gerdes, J. (2000). The Ki-67 protein: from the known and the unknown. *J Cell Physiol,* 182,(3)**,** 311-22.

Schroeter, E. H., Kisslinger, J. A. & Kopan, R. (1998). Notch-1 signalling requires ligand-induced proteolytic release of intracellular domain. *Nature,* 393,(6683)**,** 382-6.

Schubert, S., Lehmann, J., Kalfon, L., Slor, H., Falik-Zaccai, T. C. & Emmert, S. (2014). Clinical utility gene card for: Xeroderma pigmentosum. *Eur J Hum Genet,* 22,(7)**,** e1-e4.

Schwarz, T. (2005). Mechanisms of UV-induced immunosuppression. *Keio J Med,* 54,(4)**,** 165-71.

Schwarz, T. (2008). 25 years of UV-induced immunosuppression mediated by T cells-from disregarded T suppressor cells to highly respected regulatory T cells. *Photochem Photobiol,* 84,(1)**,** 10-8.

Scott, G. A., Laughlin, T. S. & Rothberg, P. G. (2014). Mutations of the TERT promoter are common in basal cell carcinoma and squamous cell carcinoma. *Mod Pathol,* 27,(4)**,** 516-23.

Sedarous, M., Keramaris, E., O'Hare, M., Melloni, E., Slack, R. S., Elce, J. S., Greer, P. A. & Park, D. S. (2003). Calpains mediate p53 activation and neuronal death evoked by DNA damage. *J Biol Chem,* 278,(28)**,** 26031-8.

Seebode, C., Lehmann, J. & Emmert, S. (2016). Photocarcinogenesis and skin cancer prevention strategies. *Anticancer Res,* 36,(3)**,** 1371-8.

Serrano, M., Hannon, G. J. & Beach, D. (1993). A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. *Nature,* 366,(6456)**,** 704-7.

Sherr, C. J. (1996). Cancer cell cycles. *Science,* 274,(5293)**,** 1672-7.

Sherr, C. J. & Roberts, J. M. (1995). Inhibitors of mammalian G1 cyclin-dependent kinases. *Genes Dev,* 9,(10)**,** 1149-63.

SIFT. [Online]. Available: www.sift.bii.a-star.edu.sg/www/SIFT_indels2.html. [Accessed 23 September 2016].

Simbolo, M., Gottardi, M., Corbo, V., Fassan, M., Mafficini, A., Malpeli, G., Lawlor, R. T. & Scarpa, A. (2013). DNA qualification workflow for next generation sequencing of histopathological samples. *PLOS One,* 8,(6)**,** 1-8.

Sinha, R. P. & Hader, D. P. (2002). UV-induced DNA damage and repair: a review. *Photochem Photobiol Sci,* 1,(4)**,** 225-36.

Slaughter, D. P., Southwick, H. W. & Smejkal, W. (1953). Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. *Cancer,* 6,(5)**,** 963-8.

Slee, E. A., O'Connor, D. J. & Lu, X. (2004). To die or not to die: how does p53 decide? *Oncogene,* 23,(16)**,** 2809-18.

Sorrell, J. M. & Caplan, A. I. (2004). Fibroblast heterogeneity: more than skin deep. *J Cell Sci,* 117,(5)**,** 667-75.

Sottoriva, A., Kang, H., Ma, Z., Graham, T. A., Salomon, M. P., Zhao, J., Marjoram, P., Siegmund, K., Press, M. F., Shibata, D. & Curtis, C. (2015). A Big Bang model of human colorectal tumor growth. *Nat Genet,* 47,(3)**,** 209-16.

Soufir, N., Moles, J. P., Vilmer, C., Moch, C., Verola, O., Rivet, J., Tesniere, A., Dubertret, L. & Basset-Seguin, N. (1999). P16 UV mutations in human skin epithelial tumors. *Oncogene,* 18,(39)**,** 5477-81.

South, A. P., Purdie, K. J., Watt, S. A., Haldenby, S., Den Breems, N. Y., Dimon, M., Arron, S. T., Kluk, M. J., Aster, J. C., McHugh, A., Xue, D. J., Dayal, J. H., Robinson, K. S., Rizvi, S. M., Proby, C. M., Harwood, C. A. & Leigh, I. M. (2014). NOTCH1 Mutations Occur Early during Cutaneous Squamous Cell Carcinogenesis. *J Invest Dermatol,* 134,(10)**,** 2630-8.

Speidel, D. (2015). The role of DNA damage responses in p53 biology. *Arch Toxicol,* 89,(4)**,** 501-17.

Spencer, J. M., Kahn, S. M., Jiang, W., DeLeo, V. A. & Weinstein, I. B. (1995). Activated ras genes occur in human actinic keratoses, premalignant precursors to squamous cell carcinomas. *Arch Dermatol,* 131,(7)**,** 796-800.

Stege, H., Roza, L., Vink, A. A., Grewe, M., Ruzicka, T., Grether-Beck, S. & Krutmann, J. (2000). Enzyme plus light therapy to repair DNA damage in ultraviolet-B-irradiated human skin. *Proc Natl Acad Sci U S A,* 97,(4)**,** 1790-5.

Stern, R. S., Weinstein, M. C. & Baker, S. G. (1986). Risk reduction for nonmelanoma skin cancer with childhood sunscreen use. *Arch Dermatol,* 122,(5)**,** 537-45.

Stone, S., Jiang, P., Dayananth, P., Tavtigian, S. V., Katcher, H., Parry, D., Peters, G. & Kamb, A. (1995). Complex structure and regulation of the P16 (MTS1) locus. *Cancer Res,* 55,(14)**,** 2988-94.

Stott, F. J., Bates, S., James, M. C., McConnell, B. B., Starborg, M., Brookes, S., Palmero, I., Ryan, K., Hara, E., Vousden, K. H. & Peters, G. (1998). The alternative product from the human CDKN2A locus, p14(ARF), participates in a regulatory feedback loop with p53 and MDM2. *EMBO J,* 17,(17)**,** 5001-14.

Stransky, N., Egloff, A. M., Tward, A. D., Kostic, A. D., Cibulskis, K., Sivachenko, A., Kryukov, G. V., Lawrence, M. S., Sougnez, C., McKenna, A., Shefler, E., Ramos, A. H., Stojanov, P., Carter, S. L., Voet, D., Cortes, M. L., Auclair, D., Berger, M. F., Saksena, G., Guiducci, C., et al. (2011). The mutational landscape of head and neck squamous cell carcinoma. *Science,* 333,(6046)**,** 1157-60.

Su, F., Viros, A., Milagre, C., Trunzer, K., Bollag, G., Spleiss, O., Reis-Filho, J. S., Kong, X., Koya, R. C., Flaherty, K. T., Chapman, P. B., Kim, M. J., Hayward, R., Martin, M., Yang, H., Wang, Q., Hilton, H., Hang, J. S., Noe, J., Lambros, M., et al. (2012). RAS mutations in cutaneous squamous-cell carcinomas in patients treated with BRAF inhibitors. *N Engl J Med,* 366,(3)**,** 207-15.

Suchniak, J. M., Baer, S. & Goldberg, L. H. (1997). High rate of malignant transformation in hyperkeratotic actinic keratoses. *J Am Acad Dermatol,* 37,(3)**,** 392-4.

Suzuki, H., Kalair, W., Shivji, G. M., Wang, B., Toto, P., Amerio, P., Kraemer, K. H. & Sauder, D. N. (2001). Impaired ultraviolet-B-induced cytokine induction in xeroderma pigmentosum fibroblasts. *J Invest Dermatol,* 117,(5)**,** 1151-5.

Svobodova, A., Walterova, D. & Vostalova, J. (2006). Ultraviolet light induced alteration to the skin. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub,* 150,(1)**,** 25-38.

Tabata, H., Nagano, T., Ray, A. J., Flanagan, N., Birch-Machin, M. A. & Rees, J. L. (1999). Low frequency of genetic change in p53 immunopositive clones in human epidermis. *J Invest Dermatol,* 113,(6)**,** 972-6.

Taguchi, M., Tsuchida, T., Ikeda, S. & Sekiya, T. (1998). Alterations of p53 gene and Ha-ras gene are independent events in solar keratosis and squamous cell carcinoma. *Pathol Int,* 48,(9)**,** 689-94.

Taguchi, M., Watanabe, S., Yashima, K., Murakami, Y., Sekiya, T. & Ikeda, S. (1994). Aberrations of the tumor suppressor p53 gene and p53 protein in solar keratosis in human skin. *J Invest Dermatol,* 103,(4)**,** 500-3.

Takata, M., Rehman, I. & Rees, J. L. (1997). p53 mutation spectrum in Japanese Bowen's disease suggests a role for mutagens other than ultraviolet light. *Int J Cancer,* 71,(3)**,** 370-2.

Takeichi, M. (1995). Morphogenetic roles of classic cadherins. *Curr Opin Cell Biol,* 7,(5)**,** 619-27.

Talghini, S., Halimi, M. & Baybordi, H. (2009). Expression of P27, Ki67 and P53 in squamous cell carcinoma, actinic keratosis and Bowen disease. *Pak J Biol Sci,* 12,(12)**,** 929-33.

Tam, K. W., Zhang, W., Soh, J., Stastny, V., Chen, M., Sun, H., Thu, K., Rios, J. J., Yang, C., Marconett, C. N., Selamat, S. A., Laird-Offringa, I. A., Taguchi, A., Hanash, S., Shames, D., Ma, X., Zhang, M. Q., Lam, W. L. & Gazdar, A. (2013). CDKN2A/p16 inactivation mechanisms and their relationship to smoke exposure and molecular features in non-small-cell lung cancer. *J Thorac Oncol,* 8,(11)**,** 1378-88.

The Human Protein Atlas. [Online]. Available: www.proteinatlas.org. [Accessed 17 September 2016].

Thompson, S. C., Jolley, D. & Marks, R. (1993). Reduction of solar keratoses by regular sunscreen use. *N Engl J Med,* 329,(16)**,** 1147-51.

Toll, A., Salgado, R., Yebenes, M., Martin-Ezquerra, G., Gilaberte, M., Baro, T., Sole, F., Alameda, F., Espinet, B. & Pujol, R. M. (2009). MYC gene numerical aberrations in actinic keratosis and cutaneous squamous cell carcinoma. *Br J Dermatol,* 161,(5)**,** 1112-8.

Toll, A., Salgado, R., Yebenes, M., Martin-Ezquerra, G., Gilaberte, M., Baro, T., Sole, F., Alameda, F., Espinet, B. & Pujol, R. M. (2010). Epidermal growth factor receptor gene numerical aberrations are frequent events in actinic keratoses and invasive cutaneous squamous cell carcinomas. *Exp Dermatol,* 19,(2)**,** 151-3.

Tong, W. M., Hande, M. P., Lansdorp, P. M. & Wang, Z. Q. (2001). DNA strand break-sensing molecule poly(ADP-Ribose) polymerase cooperates with p53 in telomere function, chromosome stability, and tumor suppression. *Mol Cell Biol,* 21,(12)**,** 4046-54.

Tsatsou, F., Trakatelli, M., Patsatsi, A., Kalokasidis, K. & Sotiriadis, D. (2012). Extrinsic aging: UV-mediated skin carcinogenesis. *Dermatoendocrinol,* 4,(3)**,** 285-97.

Udey, M. C., Von Stebut, E., Mendez, S., Sacks, D. L. & Belkaid, Y. (2001). Skin dendritic cells in murine cutaneous leishmaniasis. *Immunobiology,* 204,(5)**,** 590-4.

Uhlén, M., Bjorling, E., Agaton, C., Szigyarto, C. A., Amini, B., Andersen, E., Andersson, A. C., Angelidou, P., Asplund, A., Asplund, C., Berglund, L., Bergstrom, K., Brumer, H., Cerjan, D., Ekstrom, M., Elobeid, A., Eriksson, C., Fagerberg, L., Falk, R., Fall, J., et al. (2005). A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol Cell Proteomics,* 4,(12)**,** 1920-32.

University Of California Santa Cruz UCSC Genome Browser. [Online]. Available: www.genome.ucsc.edu/. [Accessed 07 April 2015].

Urano, Y., Asano, T., Yoshimoto, K., Iwahana, H., Kubo, Y., Kato, S., Sasaki, S., Takeuchi, N., Uchida, N., Nakanishi, H., Arase, S. & Itakura, M. (1995). Frequent p53 accumulation in the chronically sun-exposed epidermis and clonal expansion of p53 mutant cells in the epidermis adjacent to basal cell carcinoma. *J Invest Dermatol,* 104,(6)**,** 928-32.

Vallejo-Torres, L., Morris, S., Kinge, J. M., Poirier, V. & Verne, J. (2014). Measuring current and future cost of skin cancer in England. *J Public Health (Oxf),* 36,(1)**,** 140-8.

Van Allen, E. M., Wagle, N., Stojanov, P., Perrin, D. L., Cibulskis, K., Marlow, S., Jane-Valbuena, J., Friedrich, D. C., Kryukov, G., Carter, S. L., McKenna, A., Sivachenko, A., Rosenberg, M. & Kiezun, A. (2014). Whole-exome sequencing and clinical interpretation of formalin-fixed, paraffin-embedded tumor samples to guide precision cancer medicine. *Nat Med,* 20,(6)**,** 682-8.

Van Der Pols, J. C., Williams, G. M., Pandeya, N., Logan, V. & Green, A. C. (2006). Prolonged prevention of squamous cell carcinoma of the skin by regular sunscreen use. *Cancer Epidemiol Biomarkers Prev,* 15,(12)**,** 2546-8.

Van Der Riet, P., Karp, D., Farmer, E., Wei, Q., Grossman, L., Tokino, K., Ruppert, J. M. & Sidransky, D. (1994). Progression of basal cell carcinoma through loss of chromosome 9q and inactivation of a single p53 allele. *Cancer Res,* 54,(1)**,** 25-7.

Van Der Schroeff, J. G., Evers, L. M., Boot, A. J. & Bos, J. L. (1990). Ras oncogene mutations in basal cell carcinomas and squamous cell carcinomas of human skin. *J Invest Dermatol,* 94,(4)**,** 423-5.

Van Kranen, H. J. & De Gruijl, F. R. (1999). Mutations in cancer genes of UV-induced skin tumors of hairless mice. *J Epidemiol,* 9,(Suppl 6)**,** S58-65.

Van Kranen, H. J., De Gruijl, F. R., De Vries, A., Sontag, Y., Wester, P. W., Senden, H. C., Rozemuller, E. & Van Kreijl, C. F. (1995). Frequent p53 alterations but low incidence of ras mutations in UV-B-induced skin tumors of hairless mice. *Carcinogenesis,* 16,(5)**,** 1141-7.

Van Kranen, H. J., Westerman, A., Berg, R. J., Kram, N., Van Kreijl, C. F., Wester, P. W. & De Gruijl, F. R. (2005). Dose-dependent effects of UVB-induced skin carcinogenesis in hairless p53 knockout mice. *Mutat Res,* 571,(1)**,** 81-90.

Vassilev, L. T. (2007). MDM2 inhibitors for cancer therapy. *Trends Mol Med,* 13,(1)**,** 23-31.

Vigil, D., Cherfils, J., Rossman, K. L. & Der, C. J. (2010). Ras superfamily GEFs and GAPs: validated and tractable targets for cancer therapy? *Nat Rev Cancer,* 10,(12)**,** 842-57.

Vilimas, T., Mascarenhas, J., Palomero, T., Mandal, M., Buonamici, S., Meng, F., Thompson, B., Spaulding, C., Macaroun, S., Alegre, M. L., Kee, B. L., Ferrando, A., Miele, L. & Aifantis, I. (2007). Targeting the NF-kappaB signaling pathway in Notch1-induced T-cell leukemia. *Nat Med,* 13,(1)**,** 70-7.

Vogelstein, B., Lane, D. & Levine, A. J. (2000). Surfing the p53 network. *Nature,* 408,(6810)**,** 307-10.

Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A. & Kinzler, K. W. (2013). Cancer genome landscapes. *Science,* 339,(6127)**,** 1546-58.

Vousden, K. H. (2002). Activation of the p53 tumor suppressor protein. *Biochim Biophys Acta,* 1602,(1)**,** 47-59.

Vousden, K. H. & Lu, X. (2002). Live or let die: the cell's response to p53. *Nat Rev Cancer,* 2,(8)**,** 594-604.

Wang, K., Li, M. & Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res,* 38,(16)**,** 1-7.

Wang, N. J., Sanborn, Z., Arnett, K. L., Bayston, L. J., Liao, W., Proby, C. M., Leigh, I. M., Collisson, E. A., Gordon, P. B., Jakkula, L., Pennypacker, S., Zou, Y., Sharma, M., North, J. P., Vemula, S. S., Mauro, T. M., Neuhaus, I. M., Leboit, P. E., Hur, J. S., Park, K., et al. (2011). Loss-of-function mutations in Notch receptors in cutaneous and lung squamous cell carcinoma. *Proc Natl Acad Sci U S A,* 108,(43)**,** 17761-6.

Wang, Y., Zhou, X., Weinstein, E., Maryles, B., Zhang, Y., Moore, J., Gao, D., Atencio, D. P., Rosenstein, B. S., Lebwohl, M., Chen, H. D., Xiao, T. & Wei, H. (2008). p53 gene mutations in SKH-1 mouse tumors differentially induced by UVB and combined subcarcinogenic benzo[a]pyrene and UVA. *Photochem Photobiol,* 84,(2)**,** 444-9.

Weihrauch, M., Bader, M., Lehnert, G., Wittekind, C., Tannapfel, A. & Wrbitzky, R. (2002). Carcinogen-specific mutation pattern in the p53 tumour suppressor gene in UV radiation-induced basal cell carcinoma. *Int Arch Occup Environ Health,* 75,(4)**,** 272-6.

Weinberg, R. A. (1995). The retinoblastoma protein and cell cycle control. *Cell,* 81,(3)**,** 323-30.

Wellcome Trust Sanger Centre. COSMIC Catalogue of Somatic Mutations in Cancer. *v71 - 4th Nov 2014 release.* [Online]. Available: www.cancer.sanger.ac.uk/cosmic. [Accessed 30 July 2016].

White, A. C., Tran, K., Khuu, J., Dang, C., Cui, Y., Binder, S. W. & Lowry, W. E. (2011). Defining the origins of Ras/p53-mediated squamous cell carcinoma. *Proc Natl Acad Sci U S A,* 108,(18)**,** 7425-30.

Wickett, R. R. & Visscher, M. O. (2006). Structure and function of the epidermal barrier. *Am J Infect Control,* 34,(Supp 10)**,** S98-110.

Wikonkal, N. M. & Brash, D. E. (1999). Ultraviolet radiation induced signature mutations in photocarcinogenesis. *J Investig Dermatol Symp Proc,* 4,(1)**,** 6-10.

Wilke, W. W., Robinson, R. A. & Kennard, C. D. (1993). H-ras-1 gene mutations in basal cell carcinoma: automated direct sequencing of clinical specimens. *Mod Pathol,* 6,(1)**,** 15-9.

Williams, C., Ponten, F., Ahmadian, A., Ren, Z. P., Ling, G., Rollman, O., Ljung, A., Jaspers, N. G., Uhlen, M., Lundeberg, J. & Ponten, J. (1998). Clones of normal keratinocytes and a variety of simultaneously present epidermal neoplastic lesions contain a multitude of p53 gene mutations in a xeroderma pigmentosum patient. *Cancer Res,* 58,(11)**,** 2449-55.

Wohlgemuth, S., Kiel, C., Kramer, A., Serrano, L., Wittinghofer, F. & Herrmann, C. (2005). Recognizing and defining true Ras binding domains I: biochemical analysis. *J Mol Biol,* 348,(3)**,** 741-58.

Wu, X., Bayle, J. H., Olson, D. & Levine, A. J. (1993). The p53-mdm-2 autoregulatory feedback loop. *Genes Dev,* 7,(7)**,** 1126-32.

Xing, Y., Takemaru, K., Liu, J., Berndt, J. D., Zheng, J. J., Moon, R. T. & Xu, W. (2008). Crystal structure of a full-length beta-catenin. *Structure,* 16,(3)**,** 478-87.

Xu, C., Houck, J. R., Fan, W., Wang, P., Chen, Y., Upton, M., Futran, N. D., Schwartz, S. M., Zhao, L. P., Chen, C. & Mendez, E. (2008). Simultaneous isolation of DNA and RNA from the same cell population obtained by laser capture microdissection for genome and transcriptome profiling. *J Mol Diagn,* 10,(2)**,** 129-34.

Xuan, J., Yu, Y., Qing, T., Guo, L. & Shi, L. (2013). Next-generation sequencing in the clinic: promises and challenges. *Cancer Lett,* 340,(2)**,** 284-95.

Yakubu, A. & Mabogunje, O. A. (1993). Skin cancer in African albinos. *Acta Oncol,* 32,(6)**,** 621-2.

Yanofsky, V. R., Mercer, S. E. & Phelps, R. G. (2011). Histopathological variants of cutaneous squamous cell carcinoma: a review. *J Skin Cancer,* 2011**,** 210813-26.

Yates, L. R. & Campbell, P. J. (2012). Evolution of the cancer genome. *Nat Rev Genet,* 13,(11)**,** 795-806.

Younger, S. T., Kenzelmann-Broz, D., Jung, H., Attardi, L. D. & Rinn, J. L. (2015). Integrative genomic analysis reveals widespread enhancer regulation by p53 in response to DNA damage. *Nucleic Acids Res,* 43,(9)**,** 4447-62.

Yu, H. S., Liao, W. T. & Chai, C. Y. (2006). Arsenic carcinogenesis in the skin. *J Biomed Sci,* 13,(5)**,** 657-66.

Zaravinos, A., Kanellou, P. & Spandidos, D. A. (2010). Viral DNA detection and RAS mutations in actinic keratosis and nonmelanoma skin cancers. *Br J Dermatol,* 162,(2)**,** 325-31.

Zhang, H., Ping, X. L., Lee, P. K., Wu, X. L., Yao, Y. J., Zhang, M. J., Silvers, D. N., Ratner, D., Malhotra, R., Peacocke, M. & Tsou, H. C. (2001a). Role of PTCH and p53 genes in early-onset basal cell carcinoma. *Am J Pathol,* 158,(2)**,** 381-5.

Zhang, W., Lu, Q., Xie, Z. J. & Mellgren, R. L. (1997). Inhibition of the growth of WI-38 fibroblasts by benzyloxycarbonyl-Leu-Leu-Tyr diazomethyl ketone: evidence that cleavage of p53 by a calpain-like protease is necessary for G1 to S-phase transition. *Oncogene,* 14,(3)**,** 255-63.

Zhang, W., Remenyik, E., Zelterman, D., Brash, D. E. & Wikonkal, N. M. (2001b). Escaping the stem cell compartment: sustained UVB exposure allows p53-mutant keratinocytes to colonize adjacent epidermal proliferating units without incurring additional mutations. *Proc Natl Acad Sci U S A,* 98,(24)**,** 13948-53.

Zhang, Y., Xiong, Y. & Yarbrough, W. G. (1998). ARF promotes MDM2 degradation and stabilizes p53: ARF-INK4a locus deletion impairs both the Rb and p53 tumor suppression pathways. *Cell,* 92,(6)**,** 725-34.

Zhao, R., Bu, Y. C., Lee, M. H., Bode, A. M. & Dong, Z. (2016). Implications of Genetic and Epigenetic Alterations of CDKN2A (p16INK4a) in Cancer. *EBioMedicine,* 8**,** 30-9.

Ziegler, A., Jonason, A. S., Leffell, D. J., Simon, J. A., Sharma, H. W., Kimmelman, J., Remington, L., Jacks, T. & Brash, D. E. (1994). Sunburn and p53 in the onset of skin cancer. *Nature,* 372,(6508)**,** 773-6.

Ziegler, A., Leffell, D. J., Kunala, S., Sharma, H. W., Gailani, M., Simon, J. A., Halperin, A. J., Baden, H. P., Shapiro, P. E., Bale, A. E. & Brash, D. E. (1993). Mutation hotspots due to sunlight in the p53 gene of nonmelanoma skin cancers. *Proc Natl Acad Sci U S A,* 90,(9)**,** 4216-20.

Zuckerman, V., Wolyniec, K., Sionov, R. V., Haupt, S. & Haupt, Y. (2009). Tumour suppression by p53: the importance of apoptosis and cellular senescence. *J Pathol,* 219,(1)**,** 3-15.

Appendices

Table 8.1 Selected AK samples. Patient's ID represents the order of the selected samples from the database search. Abbreviations used include M=Male, F=Female. Sample in green are selected for WES analysis.

Patient ID	Year of birth	Age at biopsy	Gender	Site of the lesion	Year of biopsy
1	1915	88	M	Forehead	2003
2	1915	88	M	Forehead	2003
3	1943	61	M	Forehead	2004
4	1923	81	M	Forearm	2004
5	1932	71	M	Left ear	2003
6	1912	91	M	Scalp	2003
7	1920	83	M	Right temple	2003
8	1925	78	M	Right ear	2003
9	1937	66	F	Left forearm	2003
10	1951	52	M	Left pre-auricular region	2003
11	1923	80	F	Nose	2003
12	1954	49	M	Right Temple	2003
13	1951	54	M	Anterior side of the neck	2005
14	1941	64	M	Chest	2005
15	1930	76	M	Shoulder/chest	2006
16	1922	84	M	Scalp	2006
17	1930	76	M	Chest	2006
18	1962	44	M	Scalp	2006
19	1920	86	F	Left side neck	2006
20	1941	65	F	Scalp	2006
21	1929	65	M	Right temple	2006
22	1952	54	F	Right wrist	2006
23	1930	76	F	Left cheek	2006
24	1923	83	F	Left elbow	2006
25	1941	72	M	Right eyebrow	2013
26	1926	87	M	Scalp	2013
27	1933	80	M	Left ear	2013
28	1939	74	M	Right temple	2013
29	1932	81	M	Right scalp	2013
30	1928	85	M	Left cheek	2013
31	1940	73	M	Left ear	2013
32	1922	91	M	Nape of the neck	2013
33	1947	66	M	Nose	2013
34	1942	71	M	Right cheek	2013
35	1921	92	M	Right side of the scalp	2013
36	1920	93	M	Right temple	2013
37	1936	77	M	Left pre-auricular area	2013
38	1924	89	M	Right ear	2013
39	1931	82	F	Nose	2013
40	1949	64	M	Mid back	2013
41	1934	79	M	Left side of the neck	2013
42	1934	79	M	Dorsum of the left hand.	2013
43	1920	93	F	Right cheek	2013
44	1939	74	M	Left side of the temple	2013
45	1921	92	F	Right arm	2013
46	1944	69	M	Scalp	2013
47	1925	88	M	Left hand	2013
48	1923	90	F	Right arm	2013
49	1926	87	M	Right ear	2013
50	1932	81	M	Left lower leg	2013
51	1927	86	F	Left jaw line	2013
52	1935	78	M	Left cheek	2013
53	1945	68	M	Central upper neck	2013
54	1951	62	M	Left cheek	2013
55	1921	92	M	Left elbow	2013

Patient ID	Year of birth	Age at biopsy	Gender	Site of the lesion	Year of biopsy
56	1941	72	M	Right middle finger	2013
57	1929	84	F	Left hand	2013
58	1935	78	F	Right Upper arm	2013
59	1955	58	M	Right hand	2013
60	1919	93	F	Left Temple	2012
61	1932	81	F	Left cheek	2013
62	1948	65	F	Nape of the neck	2013
63	1943	70	F	Left lower lid	2013
64	1922	91	M	Scalp	2013
65	1931	81	M	Left inner knee	2012
66	1935	78	M	Upper left back	2013
67	1948	65	F	Right middle finger	2013
68	1930	83	F	Left middle finger	2013
69	1934	79	M	Left elbow	2013

Table 8.2 Histological features of AKs. Both dysplastic and non-dysplastic features are shown. The dysplasia grade ranged from 0 to 7+ and was divided into 3 different degree of severity. ANPIS = adnexa are not present in section.

	Dysplasia (Grade 0 to 7+) 0 = none, 1+ to 3+ = mild, 4+ to 5+ = moderate, 6+ to 7+ = severe							Non dysplastic features				
Patient ID	Nuclear atypia (None 0, Mild +, Severe ++)	Loss of polarity (Absent 0, Present +)	Mitotic activity (Absent 0, Present +)	Basal cell layer proliferation (Absent 0, Present +)	Adnexal involvement (Absent 0, Present +)	Dyskeratosis (Absent 0, Present +)	Dysplasia grade	Acanthosis (Absent 0, Present +)	Acantholysis (Absent 0, Present +)	Hyperkeratosis (Absent 0, Present +)	Solar elastosis (Absent 0, Present +)	Vacuolization (Absent 0, Present +)
1	1	1	1	1	0	0	+4	+	+	0	+	0
2	1	1	1	1	1	0	+5	0	+	0	0	0
3	1	1	1	0	1	0	+4	0	+	+	+	0
4	1	1	1	0	0	0	+3	0	+	0	+	+
5	2	1	1	1	ANPIS	0	+5	+	-	+	0	+
6	1	1	1	0	1	0	+4	+	0	+	+	0
7	1	0	0	0	1	0	+2	+	0	0	+	+
8	2	0	1	1	0	0	+4	+	0	+	+	0
9	1	0	0	0	ANPIS	0	+1	0	0	+	0	0
10	1	0	1	0	1	0	+3	0	0	0	+	+
11	1	0	1	0	1	0	+3	+	0	+	+	+
12	1	0	1	0	1	0	+3	+	0	0	+	+
13	1	0	1	1	ANPIS	0	+3	+	0	+	+	0
14	1	0	0	0	0	0	+1	0	0	0	+	+
15	2	1	1	1	0	0	+5	+	0	+	+	0
16	2	0	1	0	1	0	+4	+	0	+	0	0
17	1	0	1	1	1	0	+4	+	+	+	+	+
18	2	1	1	1	ANPIS	0	+5	+	0	+	+	+
19	1	0	0	1	1	0	+3	+	0	0	+	0
20	2	1	1	1	1	0	+6	+	0	+	+	+
21	1	1	1	0	1	0	+4	0	+	0	+	0
22	1	1	0	0	0	0	+2	0	0	+	+	0
23	1	1	1	0	0	0	+3	+	+	0	+	0
24	1	1	1	0	ANPIS	0	+3	+	0	+	0	0
25	2	1	1	1	1	0	+6	+	0	+	+	+
26	2	1	1	0	1	0	+5	+	0	+	+	0
27	1	1	0	0	0	0	+2	+	0	+	0	0
28	2	1	1	1	1	0	+6	+	0	+	+	0
29	2	1	1	1	1	0	+6	+	0	+	+	0
30	2	1	1	1	1	0	+6	+	0	+	+	+
31	1	0	1	0	1	0	+3	+	0	+	+	+
32	1	0	1	0	0	0	+2	+	0	+	0	0
33	1	0	0	0	1	0	+2	0	0	0	+	+
34	1	0	0	0	ANPIS	0	+1	0	0	+	0	0
35	2	0	1	1	0	0	+4	+	0	+	+	0
36	1	0	1	1	0	0	+3	0	0	+	+	+
37	1	1	1	0	1	0	+4	+	0	0	+	+
38	1	0	0	0	ANPIS	0	+1	0	0	+	0	0
39	1	1	1	1	1	0	+5	0	+	0	+	+
40	1	0	0	0	0	0	+1	0	+	+	+	0
41	1	0	0	0	1	0	+2	0	0	+	+	0
42	2	1	1	1	1	0	+6	+	0	+	+	0

Patient ID	Dysplasia (Grade 0 to 7+) 0 = none, 1+ to 3+ = mild, 4+ to 5+ = moderate, 6+ to 7+ = severe							Non dysplastic features				
	Nuclear atypia (None 0, Mild +, Severe ++)	Loss of polarity (Absent 0, Present +)	Mitotic activity (Absent 0, Present +)	Basal cell layer proliferation (Absent 0, Present +)	Adnexal involvement (Absent 0, Present +)	Dyskeratosis (Absent 0, Present +)	Dysplasia grade	Acanthosis (Absent 0, Present +)	Acantholysis (Absent 0, Present +)	Hyperkeratosis (Absent 0, Present +)	Solar elastosis (Absent 0, Present +)	Vacuolization (Absent 0, Present +)
43	2	1	1	1	1	0	+6	+	0	+	0	+
44	2	1	1	1	1	0	+6	+	0	+		+
45	2	1	1	1	1	0	+6	+	0	+	+	0
46	1	0	0	0	ANP IS	0	+1	0	0	+	0	0
47	1	1	1	0	1	0	+4	+	0	+	0	0
48	1	1	1	1	1	0	+5	+	0	+	+	0
49	1	1	1	0	1	0	+4	+	0	+	+	0
50	2	1	1	1	ANP IS	0	+6	+	0	0	+	0
51	1	1	1	1	1	0	+5	+	0	+	+	0
52	2	1	1	1	1	0	+6	+	0	+	0	0
53	1	1	1	0	1	0	+4	+	0	+	0	0
54	1	1	1	1	1	0	+5	+	0	+	+	0
55	1	1	1	0	0	0	+3	+	0	+	+	0
56	2	1	1	1	1	1	+7	+	0	+	+	+
57	1	1	1	0	ANP IS	0	+3	+	0	+	+	0
58	1	0	1	0	ANP IS	0	+2	+	0	+	+	0
59	1	0	1	0	1	0	+3	+	0	+	+	0
60	2	1	1	1	ANP IS	0	+5	+	0	+	+	+
61	1	1	1	1	ANP IS	0	+4	+	+	+	+	0
62	1	1	1	1	1	0	+5	+	0	+	+	0
63	1	0	1	0	1	0	+3	0	0	0	0	0
64	2	1	1	1	1	0	+6	+	+	+	+	0
65	1	1	1	1	1	0	+5	+	0	+	+	0
66	1	0	0	0	ANP IS	0	+1	+	0	+	+	0
67	1	1	1	1	ANP IS	0	+4	+	0	+	+	0
68	2	1	1	1	ANP IS	0	+5	+	0	0	+	0
69	1	1	1	1	1	0	+5	+	0	+	+	0

Table 8.3 Number of dysplastic cells in AKs. H&E images (N= 5) at (40x) from each AK section were used to count dysplastic cells within each AK lesion. A unit of section area equal to 0.056 mm² was used to ensure the consistency of the result. Mean number of dysplastic cells were counted per unit of section area (0.056 mm²) and then per total lesional area in the 5 (40x) magnification images from each AK.

Patient ID	Number of dysplastic cells per 0.056 mm²					Mean number of dysplastic cells / (0.056 mm²)	Total dysplastic area in mm²	Mean number of dysplastic cells /Total dysplastic area in mm²
	Image 1	Image 2	Image 3	Image 4	Image 5			
1	197	258	201	234	215	221	1.56	344.7
2	154	135	139	151	142	144.2	1.85	266.7
3	182	193	171	187	177	182	1.26	229.3
4	354	389	390	398	403	386.8	2.43	939.9
5	201	185	213	189	173	192.2	2.48	476.6
6	265	245	244	240	254	249.6	1.79	446.7
7	196	183	177	204	180	188	2.05	385.4
8	384	404	371	353	376	377.6	2.84	1072.3
9	57	62	42	64	52	55.4	1.06	58.724
10	50	54	46	63	69	56.4	1.94	109.4
11	198	229	220	235	227	221.8	2.68	594.4
12	56	61	44	67	53	56.2	0.96	53.9
13	220	235	213	234	245	229.4	3.74	857.9
14	32	29	25	27	31	28.8	3.17	91.2
15	323	304	311	318	326	316.4	3.41	1078.9
16	292	279	285	291	279	285.2	2.06	587.5
17	199	184	175	197	182	187.4	1.47	275.4
18	170	143	152	155	175	159	1.16	184.4
19	133	124	119	139	135	130	1.38	179.4
20	196	187	182	193	203	192.2	2.93	563.1
21	185	162	177	169	172	173	3.26	563.9
22	152	142	138	149	146	145.4	1.42	206.4
23	167	182	174	176	179	175.6	1.28	224.7
24	410	399	402	414	411	407.2	2.38	969.1
25	250	298	280	276	264	273.6	6.03	1649.8
26	344	364	355	338	363	352.8	3.59	1266.5
27	255	264	286	290	257	270.4	1.83	494.8
28	371	369	353	365	341	359.8	5.92	2130.0
29	264	297	290	248	288	277.4	3.91	1084.6
30	350	338	340	346	364	347.6	3.73	1296.5
31	134	128	123	145	132	132.4	4.72	624.9
32	89	76	45	54	64	65.6	1.23	80.6
33	109	99	105	87	93	98.6	1.47	144.9
34	56	78	67	89	92	76.4	1.39	106.1
35	234	265	251	239	247	247.2	3.27	808.3
36	196	176	173	183	187	183	2.06	376.9
37	311	302	298	308	316	307	2.47	758.3
38	90	110	89	121	107	103.4	1.36	140.6
39	78	82	66	69	72	73.4	1.19	87.3
40	50	63	59	67	55	58.8	1.93	113.5
41	153	145	141	138	149	145.2	2.18	316.5
42	159	172	152	149	170	160.4	2.46	394.6
43	137	146	138	142	152	143	1.93	275.9
44	157	168	162	171	154	162.4	1.85	300.4
45	460	439	478	416	481	454.8	2.61	1187
46	102	113	93	103	121	106.4	1.87	198.9
47	134	129	122	139	130	130.8	2.48	324.3
48	112	98	105	102	175	118.4	2.31	273.5

Patient ID	Number of dysplastic cells per 0.056 mm²					Mean number of dysplastic cells / (0.056 mm²)	Total dysplastic area in mm²	Mean number of dysplastic cells /Total dysplastic area in mm²
	Image 1	Image 2	Image 3	Image 4	Image 5			
49	207	223	199	216	231	215.2	1.91	411.0
50	255	273	282	264	278	270.4	2.58	697.6
51	309	295	299	302	287	298.4	4.72	1408.4
52	259	272	249	263	268	262.2	2.41	631.9
53	187	163	154	159	192	171	1.75	299.2
54	393	384	371	388	386	384.4	5.69	2187.2
55	154	168	172	158	165	163.4	2.59	423.2
56	235	253	242	244	261	247	3.04	750.8
57	252	266	247	264	273	260.4	3.18	828.1
58	142	160	156	148	155	152.2	3.28	499.2
59	151	146	137	142	133	141.8	3.81	540.2
60	304	298	294	293	285	294.8	3.41	1005.3
61	176	188	162	183	203	182.4	1.48	269.9
62	282	276	297	294	283	286.4	1.92	549.8
63	141	156	147	149	153	149.2	2.89	431.2
64	216	208	224	209	213	214	3.49	746.8
65	186	178	167	176	181	177.6	2.83	502.6
66	152	157	166	154	161	158	3.27	516.6
67	206	212	207	216	213	210.8	2.61	550.2
68	374	367	382	364	385	374.4	5.93	2220.2
69	302	294	289	311	299	299	3.25	971.7

Table 8.4 Number of p53 positive keratinocytes within AKs. Mean number of p53 positive cells were obtained from counting the p53 positive keratinocytes per unit of section area (0.056 mm²) in 5 (40x) magnification images from each AK. Then, the mean number of p53 positive keratinocytes within the whole lesional area was measured.

Patient ID	Number of p53 positive cells per 0.056 mm²					Mean number of p53 positive cells / (0.056 mm²)	Total lesion dysplastic area in mm²	Mean number of p53 positive cells /Total lesion dysplastic area in mm²
	Image 1	Image 2	Image 3	Image 4	Image 5			
1	58	51	74	75	102	72	1.56	112.3
2	73	82	161	116	106	107.6	1.85	199.1
3	67	85	45	89	67	70.6	1.26	88.9
4	89	109	163	167	155	136.6	2.43	331.9
5	47	53	79	59	46	56.8	2.48	140.8
6	205	235	224	220	214	219.6	1.79	393.1
7	35	47	39	51	38	42	2.05	86.1
8	354	364	322	208	176	284.6	2.84	808.2
9	40	38	15	25	32	30	1.06	31.8
10	20	34	26	43	24	29.4	1.94	57.0
11	165	217	198	105	187	174.4	2.68	467.4
12	49	55	39	56	44	48.6	0.96	46.6
13	150	114	165	103	123	131	3.74	489.9
14	0	0	0	0	0	0	3.17	0
15	299	345	276	281	308	301.2	3.41	1027.1
16	272	269	245	267	239	258.4	2.06	532.3
17	94	64	75	67	82	76.4	1.47	112.3
18	150	123	132	122	165	138.4	1.16	160.5
19	103	94	87	109	115	101.6	1.38	140.2
20	0	0	0	0	0	0	2.93	0
21	165	122	147	136	132	140.4	3.26	457.7
22	122	132	118	129	126	125.4	1.42	178.1
23	15	173	169	154	172	165	1.28	211.2
24	400	389	385	399	404	395.4	2.38	941.1
25	153	187	175	167	171	170.6	6.03	1028.7
26	284	274	255	238	273	264.8	3.59	950.6
27	223	234	253	263	234	241.4	1.83	441.7
28	0	0	0	0	0	0	5.92	0
29	61	66	53	51	59	58	3.91	226.7
30	43	0	0	0	22	32.5	3.73	121.2
31	110	103	122	114	117	113.2	4.72	534.3
32	32	22	0	0	14	23	1.23	28.3
33	109	98	99	83	87	95.2	1.47	139.9
34	0	0	0	0	0	0	1.39	0
35	23	0	0	0	14	7.3	3.27	23.8
36	120	132	127	123	112	122.8	2.06	252.9
37	233	210	206	216	211	215.2	2.47	531.5
38	45	87	64	110	59	73	1.36	99.2
39	23	25	17	29	27	24.2	1.19	28.7
40	45	55	51	57	44	50.4	1.93	97.2
41	98	89	73	91	79	86	2.18	187.4
42	157	162	102	108	169	139.6	2.46	343.4
43	112	119	109	105	116	112.2	1.93	216.5
44	146	151	143	127	136	140.6	1.85	260.1
45	456	432	459	413	441	440.2	2.61	1148.9
46	23	17	37	0	0	15.4	1.87	28.8
47	87	76	67	69	79	59.8	2.48	148.3

Patient ID	Number of p53 positive cells per 0.056 mm²					Mean number of p53 positive cells / (0.056 mm²)	Total lesion dysplastic area in mm²	Mean number of p53 positive cells /Total lesion dysplastic area in mm²
	Image 1	Image 2	Image 3	Image 4	Image 5			
48	83	98	102	93	87	92.6	2.31	213.9
49	191	184	199	173	179	178.6	1.91	341.1
50	255	271	276	260	278	268	2.58	691.4
51	112	119	98	109	114	110.4	4.72	521.1
52	209	221	213	202	219	212.8	2.41	512.8
53	45	53	41	58	56	50.4	1.75	88.2
54	87	81	79	84	67	79.6	5.69	452.9
55	74	89	84	91	77	83	2.59	214.9
56	229	241	233	242	252	239.4	3.04	727.7
57	102	111	95	104	98	102	3.18	324.3
58	64	68	73	61	59	65	3.28	213.2
59	57	47	58	45	55	52.4	3.81	199.6
60	299	294	284	279	285	288.2	3.41	982.7
61	0	0	0	0	0	0	1.48	0
62	0	0	0	0	0	0	1.92	0
63	0	0	0	0	0	0	2.89	0
64	0	0	0	0	0	0	3.49	0
65	37	32	41	39	45	38.8	2.83	109.8
66	42	57	46	45	51	48.2	3.27	157.6
67	0	0	0	0	0	0	2.61	0
68	363	342	366	351	375	359.4	5.93	2131.2
69	300	274	283	291	267	283	3.25	919.7

Table 8.5 Percentage of p53 staining in relation to the grade of dysplasia within AKs. The percentage of p53 positive cells within the lesional dysplastic tissue of AK sections was calculated by dividing the mean number of p53 positive cells over the mean number of dysplastic cells within unit of section area (0.056 mm²).

Patient ID	Dysplasia grade (1+ to 3+ = mild, 4+ to 5+ = moderate, 6+ to 7+ = severe)	Mean number of p53 positive cells / Unit of section area (0.056mm²)	Mean number of dysplastic cells / Unit of section area (0.056mm²)	Mean p53 positive cells/Mean dysplastic cells count %
1	+4	72	221	32.6%
2	+5	107.6	144.2	74.61%
3	+4	70.6	182	38.7%
4	+3	136.6	386.8	35.2%
5	+6	56.8	192.2	29.5%
6	+4	219.6	249.6	88.8%
7	+2	42	188	22.3
8	+4	284.6	377.7	75.3%
9	+1	30	55.4	54.1%
10	+3	29.4	56.4	52.1%
11	+3	174.4	221.8	78.6%
12	+3	48.6	56.2	86.4%
13	+3	131	229.4	57.1%
14	+1	0	28.8	0
15	+5	301.2	316.4	95.1%
16	+4	258.4	285.2	90.6%
17	+4	76.4	187.4	40.7%
18	+5	138.4	159	87%
19	+3	101.6	130	78.1%
20	+7	0	192.2	0%
21	+4	140.4	173	81.1%
22	+2	125.4	145.4	86.2%
23	+3	165	175.6	93.9%
24	+3	395.4	407.2	97%
25	+7	170.6	273.6	62.3%
26	+5	264.8	352.8	75.2%
27	+2	241.4	270.4	89.2%
28	+6	0	359.8	0%
29	+6	58	277.4	20.9%
30	+6	32.5	347.6	9.34%
31	+3	113.2	132.4	45.4%
32	+2	23	65.6	35%
33	+2	95.2	98.6	96.5%
34	+1	0	106.196	0
35	+4	7.3	122.8	59.4%
36	+3	122.8	183	67.1%
37	+4	215.2	307	70%
38	+1	99.28	140.624	70.6%
39	+5	24.2	73.4	32.9%
40	+1	50.4	58.8	85.7%
41	+2	86	145.2	59.2%
42	+6	139.6	160.4	87%
43	+6	112.2	143	78.4
44	+6	140.6	162.4	86.5%
45	+6	440.2	454.8	97.8%
46	+1	28.79	198.96	14.5%
47	+4	59.8	130.8	46%
48	+5	92.6	118.4	78.2%
49	+4	178.6	215.2	82.9%
50	+6	268	270.4	99%
51	+5	110.4	298.4	36.9%
52	+6	212.8	262.2	81.1%
53	+4	50.4	171	29.4%
54	+6	79.6	384.4	20.7%
55	+3	83	163.4	50.7%

Patient ID	Dysplasia grade (1+ to 3+ = mild, 4+ to 5+ = moderate, 6+ to 7+ = severe)	Mean number of p53 positive cells / Unit of section area (0.056µm²)	Mean number of dysplastic cells / Unit of section area (0.056µm²)	Mean p53 positive cells/mean dysplastic cells count %
56	+6	239.4	247	96.9%
57	+3	102	260.4	39.1%
58	+2	65	152.2	42.7%
59	+3	52.4	141.8	36.9%
60	+5	288.2	294.8	97.7%
61	+4	0	182.4	0%
62	+5	0	286.4	0%
63	+3	0	149.2	0%
64	+6	0	214	0%
65	+5	38.8	177.6	21.8
66	+1	48.2	158	30.5%
67	+4	0	210.8	0
68	+5	359.4	374.4	95.9%
69	+5	283	299	94.6%

Table 8.6 Analysis of beta catenin staining within AK. One image of (2x) magnification from each AK section was analysed. 40x images were used to examine the staining where it was not clear under the (2x). Both dysplastic lesional area and non-lesional (adjacent normal looking skin) areas were examined. The intensity of the staining within dysplastic area was scored as (+2) when it is as the normal looking skin, (+1) when it is reduced and (0) when it is completely lost. A subjective scoring to the percentage of the staining (0 -100%) was also performed. Note; when only lesional epithelium was present on the section the intensity of the staining where compared with the other AK sections included within the same experiment.

Patient ID	Dysplastic lesional area		Adjacent normally looking skin	
	Intensity of beta catenin staining	Cellular localisation	Intensity of beta catenin staining	Cellular localisation
1	+2	~100% membranous	+2	membranous
2	+1	~90% membranous	+2	membranous
	0	~10% loss of staining		
3	+1	~90% membranous	+2	membranous
	+2	~10% membranous		
4	+2	~100% membranous	+2	membranous
5	+1	~80% membranous	+2	membranous
	0	~20% Loss of staining		
6	+2	~100% membranous	+2	membranous
7	+2	~90% membranous	+2	membranous
	0	~10% loss of staining		
8	+2	~20% membranous	+2	membranous
	0	~5% loss of staining		
	+1	~70% membranous		
	+2	~5% nuclear		
9	+2	~100% membranous	+2	membranous
10	+2	~100% membranous	+2	membranous
11	+1	~100% membranous	+2	membranous
12	0	~15% loss of staining	+2	membranous
	+2	~35% membranous		
	+1	~45% membranous		
	+2	<5% nuclear		
13	0	~15% loss of staining	+2	membranous
	+2	~40% membranous		
	+1	~40% membranous		
	+2	~5% nuclear		
14	+2	~100% membranous	+2	membranous
15	+2	~15% membranous	+2	membranous
	+1	~85% membranous		
16	0	~5% loss of staining	+2	membranous
	+2	~45% membranous		
	+1	~50% membranous		
17	+2	~5% nuclear	+2	membranous
	+1	~95% membranous		
18	0	~20% loss of staining	Only lesional epithelium	
	+1	~80% membranous		
19	+1	~100% loss of staining	+2	membranous
20	+2	~10% membranous	+2	membranous
	+1	~90% membranous		
21	0	~10% loss of staining	+2	membranous
	+1	~80% membranous		
22	0	~100% loss of staining	Only lesional epithelium	
23	+1	~100% membranous	Only lesional epithelium	
24	+1	~100% membranous	+2	membranous
25	+1	~100% membranous	+2	membranous
26	+1	~85% membranous	Only lesional epithelium	
	+2	~10% membranous		

Patient ID	Dysplastic lesional area		Adjacent normally looking skin	
	Intensity of beta catenin staining	Cellular localisation	Intensity of beta catenin staining	Cellular localisation
26	+1	~5% nuclear		
27	+2	~10% membranous	Only lesional epithelium	
	+1	~80% membranous		
	+2	~10% cytoplasmic		
28	+2	~10% membranous	+2	membranous
	+1	~90% membranous		
29	+1	~80% membranous	+2	membranous
30	+2	~30% membranous	+2	membranous
	+1	~30% membranous		
	0	~40% loss of staining		
31	+2	~30% membranous	+2	membranous
	+1	~70% membranous		
32	+1	~100% membranous	Only lesional epithelium	
33	+2	~100% membranous	+2	membranous
34	+2	~100% membranous	+2	membranous
35	+2	~50% membranous	+2	membranous
	0	~50% loss of staining		
36	+1	~15% membranous	+1	membranous
	0	~85% loss of staining		
37	+2	~80% membranous	+2	membranous
	+1	~20% membranous		
38	+1	~100% membranous	+2	membranous
39	+1	~50% membranous		
	+2	~50% membranous		
40	0	~80% loss of staining	+2	membranous
	+1	~20% membranous		
41	+1	~70% membranous	+2	membranous
	+2	~30% membranous		
42	+1	~65% membranous	+2	membranous
	0	~35% loss of staining		
43	+1	~100% membranous	Only lesional epithelium	
44	+1	~80% membranous	+2	membranous
	0	~20% loss of staining		
45	+2	~60% membranous	+2	membranous
	+1	~40% membranous		
46	0	100% loss of staining	Only lesional epithelium	
47	+2	~40% membranous	Only lesional epithelium	
	+1	~60% membranous		
48	+2	~80% membranous	+2	membranous
	+1	~10% membranous		
	+2	~10% cytoplasmic		
49	+1	~100% membranous	Only lesional epithelium	
50	+1	~95% membranous	+2	membranous
	+2	~5% nuclear		
51	+1	~95% membranous	+2	membranous
	+2	~5% nuclear		
52	+1	~90% membranous	+2	membranous
	+2	~10% cytoplasmic		
53	+1	~90% membranous	Only lesional epithelium	
	+2	~5% nuclear		
	+2	~5% cytoplasmic		
54	+1	~60% membranous	+2	membranous
	0	~30% loss of staining		
	+2	~10% nuclear		
55	+1	~90% membranous	+2	membranous
	+2	~10% membranous		

Patient ID	Dysplastic lesional area		Adjacent normally looking skin	
	Intensity of beta catenin staining	Cellular localisation	Intensity of beta catenin staining	Cellular localisation
56	+2	~80% membranous	+2	membranous
	+1	~20% membranous		
57	+2	~80% membranous	+2	membranous
	+1	~15% membranous		
	+2	~5% nuclear		
58	+1	~100% membranous	+2	membranous
48	+2	~80% membranous	+2	membranous
	+1	~10% membranous		
	+2	~10% cytoplasmic		
49	+1	~100% membranous	Only lesional epithelium	
50	+1	~95% membranous	+2	membranous
	+2	~5% nuclear		
51	+1	~95% membranous	+2	membranous
	+2	~5% nuclear		
52	+1	~90% membranous	+2	membranous
	+2	~10% cytoplasmic		
53	+1	~90% membranous	Only lesional epithelium	
	+2	~5% nuclear		
	+2	~5% cytoplasmic		
54	+1	~60% membranous	+2	membranous
	0	~30% loss of staining		
	+2	~10% nuclear		
55	+1	~90% membranous	+2	membranous
	+2	~10% membranous		
56	+2	~80% membranous	+2	membranous
	+1	~20% membranous		
57	+2	~80% membranous	+2	membranous
	+1	~15% membranous		
	+2	~5% nuclear		
58	+1	~100% membranous	+2	membranous
59	+2	~100% membranous		membranous
60	+2	~90% membranous	+2	membranous
	+1	~10% membranous		
61	+2	~100% membranous	Only lesional epithelium	
62	+1	~90% membranous	+2	membranous
	+2	~10% nuclear		
63	+2	~95% membranous	+2	membranous
	+1	~5% membranous		
64	+2	~40% membranous	+2	membranous
	+1	~60% membranous		
65	+2	~20% membranous	Only lesional epithelium	
	+1	~70% membranous		
	+2	~10% nuclear		
66	+2	~60% membranous	+2	membranous
	+1	~40% membranous		
67	+1	~10% membranous	Only lesional epithelium	
	+2	~90% membranous		
68	+2	~100% membranous	Only lesional epithelium	
69	+2	~100% membranous	+2	membranous

Table 8.7 Counting the number of CD4 positive cells within AK sections. The mean number of CD4 positive cells was counted per unit of section area (0.056 mm^2) in 5 (40x) images from each section. The percentage of the total CD4 positive immune cells per total perilesional infiltrate was also counted.

Patient ID	Number of CD4 positive cells / Number of perilesional immune infiltrate					Mean CD4 positive cells/ Mean perilesional immune infiltrate	%CD4 positive cells/Perilesional immune infiltrate
	Image 1	Image 2	Image 3	Image 4	Image 5		
1	24/135	29/153	22/151	35/148	18/66	25.6/130.6	19.6
2	45/168	37/154	43/161	61/143	44/156	46/156.4	29.4
3	10/99	17/87	13/77	21/107	11/74	14.4/88.8	16.2
4	206/454	188/432	177/465	196/442	203/468	194/452.2	42.9
5	33/184	18/170	28/177	31/149	21/166	26.2/169.2	15.5
6	39/55	27/44	9/65	15/55	27/57	23.4/55.2	42.4
7	38/144	49/165	32/177	38/165	55/166	42.4/163.4	25.9
8	11/33	18/46	28/51	16/54	25/61	19.6/49	40
9	0/27	0/41	0/35	0/19	0/25	0/28.2	0
10	44/165	49/168	56/134	33/187	35/155	43.4/161.8	26.8
11	155/243	133/235	147/282	137/261	145/278	143.4/259.8	55.2
12	27/66	28/91	38/78	38/99	25/87	31.2/84.2	37.1
13	44/133	32/142	52/156	55/141	46/138	45.8/142	32.3
14	4/79	16/71	11/65	18/59	6/61	11/67	16.4
15	33/67	54/71	43/79	56/71	49/89	47/75.4	62.3
16	55/135	49/165	53/118	49/173	38/148	48.8/147.8	33
17	16/103	22/95	5/71	25/114	12/67	16/90	17.8
18	147/233	144/248	132/219	105/241	139/253	133.4/238.8	55.9
19	26/94	35/98	30/77	7/61	24/77	24.4/81.4	29.9
20	114/274	125/265	126/285	106/235	127/281	119.6/268	44.6
21	52/124	67/125	62/134	43/115	92/165	63.2/132.6	47.7
22	35/88	21/96	39/105	38/92	23/131	31.2/102.4	30.5
23	84/142	73/124	62/115	71/128	77/135	73.4/128.8	56.9
24	68/314	96/323	79/351	85/351	88/358	83.2/339.4	24.5
25	143/268	112/287	119/234	154/216	155/285	136.6/258	52.9
26	0/315	0/348	0/398	0/412	0/435	0/381.6	0
27	31/76	45/135	18/96	18/82	9/91	24.2/96	25.2
28	25/125	18/132	35/156	25/154	21/152	24.8/143.8	17.3
29	114/206	124/234	105/232	92/267	99/208	106.8/229.4	46.6
30	65/79	22/68	33/61	19/53	17/65	31.2/65.2	47.9
31	35/65	25/86	38/74	17/52	35/68	29.4/69	42.6
32	102/215	124/267	116/245	92/235	94/205	105.6/233.4	45.2
33	25/54	16/49	16/62	28/51	13/49	19.6/53	36.9
34	45/98	51/69	35/92	65/82	33/74	45.8/83	55.9
35	15/45	13/49	19/56	12/46	19/65	15.6/52.2	29.8
36	66/118	74/135	68/127	54/135	48/115	62/126	49.2
37	0/64	0/65	0/69	0/67	0/63	0/65.6	0
38	35/65	23/58	33/67	18/46	31/56	28/58.4	47.9
39	43/81	32/64	38/76	34/86	23/67	34/74.8	45.5
40	0/32	0/35	0/33	0/34	0/36	0/34	0
41	96/131	102/157	119/184	124/153	91/124	105.8/149.8	70.6
42	93/189	85/207	73/179	82/174	62/165	79/182.8	43.2
43	0/45	0/53	0/36	0/39	0/51	0/44.8	0
44	88/322	74/345	83/340	79/312	81/353	81/334.4	24.2
45	35/79	31/68	28/65	29/65	20/75	28.6/70.4	40.6
46	0/21	0/51	0/38	0/35	0/26	0/34.2	0
47	0/36	0/44	0/46	0/32	0/34	0/38.4	0
48	16/66	21/65	10/45	16/71	18/62	16.2/61.8	26.2
49	69/189	65/182	56/184	79/192	71/187	68/186.8	36.4
50	63/126	75/118	24/137	39/96	41/92	48.4/113.8	42.5
51	77/156	63/182	65/164	59/206	65/182	65.8/178	36.9
52	13/65	18/52	19/47	29/63	24/59	20.6/57.2	36

Patient ID	Number of CD4 positive cells / Number of perilesional immune infiltrate					Mean CD4 positive cells/ Mean perilesional immune infiltrate	%CD4 positive cells/Perilesional immune infiltrate
	Image 1	Image 2	Image 3	Image 4	Image 5		
53	41/135	45/118	54/112	65/157	65/109	54/126.2	42.8
54	63/172	81/169	82/154	65/156	65/156	71.2/161.4	44.1
55	32/115	23/151	65/119	36/84	25/78	36.2/133.4	27.1
56	43/125	54/131	68/108	56/113	51/91	54.4/113.6	47.8
57	63/153	67/168	68/144	61/145	74/152	66.6/152.4	43.7
58	16/54	31/56	29/61	18/54	29/56	24.6/56.2	43.8
59	65/113	65/115	35/98	65/124	65/118	59/113.6	51.9
60	11/36	8/46	9/24	15/32	16/41	11.8/35.8	32.9
61	22/246	16/243	28/252	21/244	25/253	22.4/247.6	9
62	0/119	0/129	0/121	0/107	0/196	0/134.4	0
63	22/95	19/76	32/85	16/86	23/72	22.4/82.8	27
64	115/245	116/237	107/265	89/203	159/251	117.2/240.2	48.8
65	68/254	123/285	68/234	106/287	106/268	94.2/265.6	35.4
66	0/58	0/53	0/59	0/54	0/57	0/56.2	0
67	52/86	65/84	58/86	76/94	65/92	63.2/88.4	71.5
68	105/185	134/204	98/187	96/193	112/198	109/193.4	56.4
69	86/208	95/168	84/178	62/186	95/178	84.4/183.6	45.9

Table 8.8 Counting the number of CD8 positive cells within the AK samples. The mean number of CD8 positive cells was counted per unit of section area (0.056 mm²) within 5 (40x) images from each section. The percentage of the total CD8 positive immune cells per total perilesional infiltrate was also counted.

Patient ID	Number of CD8 positive cells / Number of perilesional immune infiltrate					Mean CD8 positive cells/ Mean perilesional immune infiltrate	%CD8 positive cells/Perilesional immune infiltrate
	Image 1	Image 2	Image 3	Image 4	Image 5		
1	32/154	47/149	44/167	36/172	57/169	58/128	45.3
2	44/163	36/152	49/167	51/132	33/146	42.6/152	28
3	34/57	22/48	7/43	14/54	26/58	20.6/52	39.6
4	204/464	198/458	187/438	195/478	209/453	198.6/458.2	43.3
5	34/191	16/165	25/174	21/159	27/168	24.6/171.4	14.3
6	8/83	15/93	11/74	18/94	9/72	14.6/83.2	17.5
7	25/145	28/156	21/161	31/144	19/67	24.8/134.6	18.4
8	34/67	45/79	39/74	51/69	43/81	42.4/74	57.3
9	4/25	6/31	3/21	8/17	6/11	5.4/21	25.7
10	46/143	37/132	49/155	51/139	44/161	45.4/146	31.1
11	134/256	123/274	117/262	107/251	115/259	119.2/260.4	45.7
12	28/67	25/87	33/96	37/89	29/77	30.4/83.2	36.5
13	48/175	43/167	59/149	32/161	33/153	43/161	26.7
14	10/35	15/43	24/47	18/37	33/51	20/42.6	47
15	23/87	29/102	37/96	31/85	25/110	29/96	30.2
16	45/165	49/147	57/129	42/161	33/143	45.2/149	30.9
17	13/95	19/86	9/78	23/104	11/90	15/90.6	16.5
18	104/199	112/214	109/223	89/207	85/197	99.8/208	47.9
19	5/67	11/62	7/54	13/59	3/51	7.8/58.6	13.3
20	134/254	123/278	117/212	127/201	145/264	129.2/241.8	53.4
21	42/134	57/129	52/124	46/125	72/155	53.8/133.4	40.3
22	72/136	64/113	58/108	63/119	76/121	66.6/119.4	55.7
23	22/116	18/121	34/127	23/124	17/132	22.8/124	18.4
24	16/332	20/356	18/405	29/389	26/420	21/380.4	5.5
25	127/233	124/248	112/219	85/241	119/253	113.4/238.8	47.4
26	78/332	84/335	93/348	71/322	75/347	80.2/336.8	23.8
27	34/107	21/114	47/93	26/86	22/74	30/94.8	31.6
28	93/129	107/148	112/164	101/143	99/139	102.4/144.6	70.8
29	96/203	108/245	103/233	94/217	89/206	98/220.8	44.3
30	29/95	35/109	11/78	17/85	7/81	19.8/89.6	22.1
31	24/81	29/88	17/67	6/59	19/76	19/74.2	25.6
32	108/261	111/282	104/277	99/256	119/274	108.2/270	40.1
33	23/54	24/58	32/69	16/40	22/45	23.4/53.2	43.9
34	34/117	31/124	57/83	27/76	32/84	36.2/96.8	37.4
35	33/64	27/56	23/59	17/48	14/55	22.8/56.4	40.4
36	67/128	76/123	59/116	52/134	64/121	63.6/124.4	51.9
37	13/44	21/53	23/41	17/51	26/45	20/46.8	42.7
38	39/95	33/48	36/57	28/66	34/66	34/66.4	51.2
39	46/79	33/68	37/74	39/81	26/69	36.2/74.2	48.7
40	10/34	11/41	15/43	9/37	16/45	12.2/40	30.5
41	57/198	66/183	49/186	62/194	54/191	57.6/190.4	30.2
42	67/175	62/183	58/174	49/189	55/191	58.2/182.4	32
43	198/262	211/272	198/217	203/266	219/254	205.8/254.2	80.6
44	20/352	29/367	34/376	27/369	22/356	26.4/364	7.2
45	34/78	39/67	23/47	28/53	19/65	28.6/62	46.1
46	0/41	0/31	0/48	0/55	0/16	0/38.2	0.00
47	34/27	27/61	22/56	31/29	39/42	30.6/43	69.8
48	43/96	53/78	39/85	57/91	32/74	44.8/84.8	52.8
49	92/199	98/214	77/189	87/189	65/184	83.8/195	42.9
50	64/114	60/103	51/108	37/94	45/96	51.4/103	49.9
51	57/143	71/138	63/149	31/105	84/152	61.2/137.4	44.5
52	14/65	19/52	9/40	11/61	16/58	13.8/55.2	25

Patient ID	Number of CD8 positive cells / Number of perilesional immune infiltrate					Mean CD8 positive cells/ Mean perilesional immune infiltrate	%CD8 positive cells/Perilesional immune infiltrate
	Image 1	Image 2	Image 3	Image 4	Image 5		
53	99/178	111/193	91/179	85/189	112/184	99.6/184.6	53.9
54	113/198	102/178	98/164	73/156	105/168	98.2/172.8	56.8
55	42/76	56/72	49/78	67/83	52/69	53.2/75.6	70.3
56	46/115	49/101	53/94	58/107	41/89	49.4/101.2	48.8
57	61/163	57/158	48/164	41/155	54/148	52.2/157.6	33.1
58	23/58	15/47	17/64	27/54	19/59	20.2/56.4	35.8
59	21/98	28/78	31/84	14/88	19/74	22.6/84.4	26.7
60	6/19	7/15	10/23	2/14	8/26	6.6/19.4	34
61	0/195	0/219	0/221	0/207	0/194	0/207.2	0.00
62	36/125	39/112	43/106	58/113	47/99	44.6/111	40.1
63	12/45	10/42	9/37	19/53	14/49	12.8/45.2	28.3
64	105/235	109/217	111/245	78/187	139/232	108.4/223.2	48.6
65	103/212	106/223	95/218	83/216	109/226	99.2/219	45.2
66	64/104	70/110	81/116	72/114	75/121	72.4/80.4	90
67	0/43	0/53	0/59	0/49	0/57	0/52.2	0.00
68	52/116	51/112	45/108	57/103	42/119	49.4/111.6	44.2
69	67/163	73/158	78/149	51/155	84/142	70.6/153.4	46

Table 8.9 FOXP3 positive cells count within the dermal immune infiltrate of AKs. The number of FOXP3 positive cells was counted per unit of section area (0.056 mm²) in 5 (40x) images from each section. The percentage of the total FOXP3 positive immune cells per total perilesional infiltrate was also counted.

Patient ID	Number of FOXP3 positive cells / Number of perilesional immune infiltrate					Mean FOXP3 positive cells/ Mean perilesional immune infiltrate	%FOXP3 positive cells/Perilesional immune infiltrate
	Image 1	Image 2	Image 3	Image 4	Image 5		
1	11/98	17/110	5/120	18/111	14/128	13/113.4	11.5
2	36/169	31/142	27/159	37/158	22/154	30.6/156.4	19.5
3	12/78	19/63	11/71	6/53	13/59	12.2/64.8	18.8
4	4/416	2/419	0/429	0/412	6/415	2.4/418.2	0.6
5	16/164	11/159	13/145	8/166	19/151	13.4/157	8.5
6	21/78	13/65	17/58	23/66	9/53	16.6/64	25.9
7	25/145	28/156	21/161	31/144	19/67	24.8/134.6	18.4
8	16/56	18/67	25/76	12/48	10/56	16.2/60.6	26.7
9	0/25	0/31	0/21	0/17	0/11	0/21	0
10	59/123	46/169	38/129	47/152	45/148	47/144.2	32.5
11	38/237	31/229	37/298	33/202	41/293	36/271.8	13.2
12	9/73	15/67	3/74	0/59	17/76	8.8/69.8	12.6
13	50/165	55/136	59/146	47/168	60/134	53.6/148	36.2
14	7/29	4/23	11/34	9/17	16/38	9.4/28.2	33.3
15	21/43	14/37	18/53	24/56	10/37	17.4/45.2	38.4
16	56/156	54/149	65/165	69/163	79/146	64.6/155.8	41.4
17	27/56	23/64	18/71	22/61	13/49	20/60.2	33.2
18	25/272	29/268	31/257	21/237	16/243	24.4/255.4	9.5
19	16/58	19/62	7/42	24/61	2/9	13.6/46.4	29.3
20	12/254	23/269	31/298	27/286	37/274	26/276.2	9.4
21	0/131	0/143	0/121	0/156	0/137	0/137.6	0
22	32/125	37/102	41/97	46/88	43/135	39.8/109.4	36.3
23	38/132	34/142	49/139	51/145	31/129	40.6/137.4	29.5
24	0/305	0/332	0/357	0/361	0/319	0/334.8	0
25	74/265	98/249	117/254	104/248	92/262	97/255.6	37.9
26	113/323	102/339	98/355	89/331	89/397	98.2/329	29.8
27	22/90	38/112	12/82	15/89	10/91	19.4/92.8	20.9
28	34/129	59/148	61/164	52/143	49/139	51/144.6	35.2
29	66/210	89/256	94/235	87/248	69/238	81/217.4	37.2
30	38/110	29/124	43/96	22/64	26/78	31.6/94.4	33.4
31	22/63	28/65	33/45	15/46	16/44	22.8/52.6	43.3
32	39/226	35/242	53/222	45/228	41/218	42.6/227.2	18.6
33	22/68	35/75	14/56	35/48	25/54	26.2/60.2	43.5
34	12/76	16/81	21/92	11/83	19/79	15.8/82.2	19.2
35	24/88	69/98	15/76	12/95	23/58	28.6/83	34.4
36	22/153	14/152	16/133	14/134	9/141	15/142.6	10.5
37	24/85	16/56	21/79	18/92	17/58	19.2/74	25.9
38	13/57	11/62	4/58	8/49	16/58	10.4/56.8	18.3
39	21/54	11/42	16/51	28/66	35/75	22.2/57.6	38.5
40	23/65	31/74	23/69	15/54	15/52	21.4/62.8	34.1
41	46/144	57/156	34/132	45/167	45/145	45.4/148.8	30.5
42	71/125	66/175	51/165	51/176	66/201	61/168.4	36.2
43	75/289	86/254	81/230	85/254	72/256	79.8/256.6	31.1
44	0/336	0/256	0/421	0/325	0/322	0/332	0.00
45	8/85	11/79	12/89	19/103	13/16	12.6/74.4	16.9
46	0/32	0/39	0/28	0/37	0/29	0/33	0
47	0/32	0/69	0/21	0/10	0/56	0/37.6	0
48	45/102	35/98	21/98	32/105	11/86	26.8/97.8	27.4
49	82/186	78/242	84/211	65/145	69/198	75.6/196.4	38.4
50	29/72	26/88	28/55	16/75	18/71	23.4/72.2	32.4

Patient ID	Number of FOXP3 positive cells / Number of perilesional immune infiltrate					Mean FOXP3 positive cells/ Mean perilesional immune infiltrate	%FOXP3 positive cells/Perilesional immune infiltrate
	Image 1	Image 2	Image 3	Image 4	Image 5		
51	82/213	78/163	85/179	45/166	70/186	65.6/181.4	36.1
52	29/88	32/96	31/69	23/85	37/94	30.4/86.4	35.1
53	69/178	81/193	61/179	55/189	82/184	69.6/184.6	37.8
54	45/135	61/165	46/158	25/115	67/198	48.8/154.2	31.6
55	22/105	24/95	11/84	19/108	25/92	20.2/96.8	20.8
56	33/99	57/95	18/109	22/95	23/104	30.6/100.4	30.4
57	45/145	34/133	39/125	55/140	42/114	43/131.4	32.7
58	9/38	12/49	6/28	11/37	10/48	9.6/40	24
59	7/71	49/164	11/73	0/69	16/77	16.6/90.8	18.2
60	17/40	9/38	13/45	18/54	14/35	14.2/42.4	33.4
61	40/234	45/254	47/221	39/209	54/219	45/227.4	19.7
62	12/142	15/141	6/135	12/142	15/135	12/139	8.6
63	0/76	0/88	0/93	0/99	0/76	0/86.4	0
64	19/254	16/258	17/248	12/263	21/253	17/255.2	7.5
65	43/215	42/228	51/218	65/221	53/205	50.8/217.4	23.3
66	17/58	12/66	22/58	16/42	18/53	17/55.4	30.6
67	0/55	0/65	0/43	0/61	0/57	0/56.2	0
68	35/123	34/157	35/117	48/98	36/123	37.6/123.6	30.4
69	26/115	38/96	25/115	27/128	35/131	30.2/117	25.8

Table 8.10 First targeted sequencing panel gene list. Base pair alterations were previously reported within the listed genes in either cSCC, AK, Bowen's disease lesions or within all lesions together. [1] indicates mutations recognised using whole exome sequencing, while [2] indicates mutations identified using PCR and Sanger sequencing

Gene	Gene function	The lesion	The study
SYNE2	Encodes nuclear outer membrane protein which binds cytoplasmic F-actin	cSCC[1]	South et al., 2014
KALRN	Encodes protein that interacts with the huntingtin-associated protein 1 and play a role in nerve and axonal development	cSCC[1]	South et al., 2014
MUC5B	Encodes protein that is important in mucin production	cSCC[1]	South et al., 2014
TPO	Encodes a membrane-bound glycoprotein that catalyses the iodination of tyrosine molecule within thyroglobulin complex.	cSCC[1]	South et al., 2014
PKHD1	Encodes receptor protein that involved in tubulogenesis and keeping patent ducts within epithelium.	cSCC[1]	South et al., 2014
APOB	A major apolipoprotein of chylomicrons and low density lipoproteins. Loss of gene function is associated with disturbance in plasma lipid levels.	cSCC[1]	South et al., 2014
DNAH5	Belongs to Dyneins superfamily (microtubule-associated motor protein complexes). DNAH2 is an axonemal inner arm dynein heavy chain and located in cilia and flagella.	cSCC[1]	South et al., 2014
FSIP2	Fibrous sheath interacting protein2. It is important in formation of the cytoskeletal of sperm flagellum.	cSCC[1]	South et al., 2014
SRRM2	Act as splicing co-activator for serine receptor family protein.	cSCC[1]	South et al., 2014
PRB2	Basic salivary proline-rich protein 2	cSCC[1]	South et al., 2014
SPTA1	Maintaining the integrity of the cell membrane of blood cells	cSCC[1]	South et al., 2014
XIRP2	Preventing the depolymerisation of actin filaments.	cSCC[1]	South et al., 2014
SIGLEC1	Member of the immunoglobulin superfamily that mediates clathrin dependent endocytosis	cSCC[1]	South et al., 2014
CNTNAP5	Control the development and function of the nervous system.	cSCC[1]	South et al., 2014
ABCA13	A member of ATP-binding cassette (ABC) family. Act as transmembrane transporter protein.	cSCC[1]	South et al., 2014
CDH23	A member of the cadherin superfamily (important for normal hearing).	cSCC[1]	South et al., 2014
CSMD1	Tumour suppressor gene. Supress the development of squamous cell carcinomas.	cSCC[1]	South et al., 2014
UNC80	Component of sodium channel complex. Important in stimulation of neurons.	cSCC[1]	South et al., 2014
OBSCN	Have a role in mediating the interactions between the sarcoplasmic reticulum and myofibrils	cSCC[1]	South et al., 2014
CSMD3	It has different expression in different tissues.	cSCC[1]	South et al., 2014
RYR2	Important in regulation calcium levels within cardiac muscle fibre	cSCC[1]	South et al., 2014
GPR98	A member of the G-protein receptor family protein and expressed in the central nervous system	cSCC[1]	South et al., 2014
NOTCH2	Controlling cell fate decisions/ Tissue specific and time specific functions	cSCC[1]	South et al., 2014
MUC16	Expressed on eye with un-determined function.	cSCC[1]	South et al., 2014
DNAH5	Has ATPase activity with a role in energy generation	cSCC[1]	South et al., 2014
LRP1B	It interact with different ligands (involved in many cell functions)	cSCC[1]	South et al., 2014
TTN	Encodes a large abundant protein of striated muscle.	cSCC[1]	South et al., 2014
TP53	This gene encodes a tumour suppressor protein that expressed in response to diverse cellular stresses to regulate apoptosis, senescence and cell cycle arrest.	cSCC[1], AK[2], Bowen's disease[2]	[South et al., 2014 (cSCC)] & [Park et al., 1996, Taguchi et al., 1998, Ziegler et al., 1994 (AK)] & [Campbell et al., 1993b, Takata et al., 1997 (BD)]
PCLO	Important in starting synaptic activity and in transferring synaptic vesicle	cSCC[1]	South et al., 2014
CACNA1C	A part of voltage-dependent calcium channel mediates CA2+ influx	cSCC[1]	South et al., 2014
MUC4	Encodes glycoproteins with lubrication function.	cSCC[1]	South et al., 2014
DNAH17	It has association with axonemal dynein	cSCC[1]	South et al., 2014

Gene	Gene function	The lesion	The study
FLG	Important in aggregation of keratin intermediate filaments in mammalian epidermis	cSCC[1]	South et al., 2014
SYNE1	Expressed mainly in skeletal and smooth muscle.	cSCC[1]	South et al., 2014
USH2A	Localised within basement membranes and is crucial for inner ear and retinal development.	cSCC[1]	South et al., 2014
CDKN2A	Tumour suppressor gene, encodes (INK4A and ARF) genes, involved mainly in control cell cycle at G1 and p53 stabilisation	cSCC[1] and AK[2] cSCC[2]	[South et al., 2014 (cSCC)] & [Kanellou et al., 2008, Saridaki et al., 2003, Soufir et al., 1999 (cSCC & AK)].
CNTNAP2	Encodes cell adhesion and receptor protein within nervous system.	cSCC[1]	South et al., 2014
DNAH3	It has association with axonemal dynein heavy chain	cSCC[1]	South et al., 2014
FAT1	A member of the cadherin superfamily that act as adhesion molecules and/or signalling receptors.	cSCC[1]	South et al., 2014
GPR158	A member of the G-protein receptor family protein	cSCC[1]	South et al., 2014
HRNR	Information not available	cSCC[1]	South et al., 2014
KIAA1211	Information not available	cSCC[1]	South et al., 2014
MUC17	Encodes glycoproteins with lubrication function.	cSCC[1]	South et al., 2014
NEB	Express a giant protein within the cytoskeletal matrix.	cSCC[1]	South et al., 2014
PAPPA2	Encodes one of metalloproteinases and is important in regulation of insulin level within the cells.	cSCC[1]	South et al., 2014
RELN	Regulate cell-cell interactions of nerve cells.	cSCC[1]	South et al., 2014
TEX15	Information not available	cSCC[1]	South et al., 2014
ACACB	A part of a complex multifunctional enzyme system. Important in fatty acid oxidation.	cSCC[1]	South et al., 2014
CACNA1B	Important in the release of neurotransmitter from nerve cells	cSCC[1]	South et al., 2014
COL6A3	Involved in the formation of type VI collagen.	cSCC[1]	South et al., 2014
DNAH10	It has association with axonemal dynein heavy chain	cSCC[1]	South et al., 2014
DNAH8	It has association with axonemal dynein heavy chain (sperm and cilia of respiratory system).	cSCC[1]	South et al., 2014
DST	A part of adhesion junction plaque proteins.	cSCC[1]	South et al., 2014
FBN1	Encodes proteins that involved in maintaining and supporting the elasticity of connective tissue.	cSCC[1]	South et al., 2014
MAP1A	Belongs to the microtubule-associated protein family. Play a role in the process of neurogenesis.	cSCC[1]	South et al., 2014
MLL3	Information not available	cSCC[1]	South et al., 2014
PCDH44	Information not available	cSCC[1]	South et al., 2014
RYR3	Important in the release of calcium from intracellular sources.	cSCC[1]	South et al., 2014
ZFHX4	it may has a role in neural and muscle differentiation	cSCC[1]	South et al., 2014
DMXL2	Encodes multifunction protein that involved in cellular signal transduction.	cSCC[1]	South et al., 2014
DNAH7	Form a part of inner dynein arm of ciliary axonemes	cSCC[1]	South et al., 2014
FLNB	A member of the filamin family. Important in repairing vascular damages	cSCC[1]	South et al., 2014
HMCN1	Encodes a member of the immunoglobulin superfamily and connect the mechanosensory neurons to the epidermis.	cSCC[1]	South et al., 2014
MLL2	Information not available	cSCC[1]	South et al., 2014
NOTCH1	Control different developmental processes and control the decision of cell fate.	cSCC[1]	South et al., 2014
SORCS3	Belongs to vacuolar protein sorting 10 (VPS10) domain-containing receptor proteins superfamily	cSCC[1]	South et al., 2014
C1orf173	Information not available	cSCC[1]	South et al., 2014
WT1	Important in normal development of urogenital system	cSCC[1]	Durinck et al., 2011
EP300	Function as histone acetyltransferase. Regulate cell proliferation and differentiation through controlling chromatin remodelling.	cSCC[1]	Durinck et al., 2011
PIK3CG	Control the structure and function of epithelium through modulation of extra-cellular signals	cSCC[1]	Durinck et al., 2011

Gene	Gene function	The lesion	The study
EZH2	Tumour suppresser gene maintains genes to be transcriptionally inactive over consecutive cellular generations	cSCC[1]	Durinck et al., 2011
HRAS	Oncogene involved in signal transduction pathways	cSCC[1]	Durinck et al., 2011
HSPB2	Maintaining muscle structure and function	cSCC[1]	Durinck et al., 2011
TERT	Involved in telomeres shortening, cellular senescence	cSCC[2] and Bowen's disease[2]	(Griewank et al., 2013, Scott et al., 2014)
EGFER	A member of protein kinase family involved in cell proliferation	cSCC[2]	Mauerer et al., 2011
PTCH1	Tumour suppressor gene. Expressed protein forms a receptor for sonic hedgehog protein which is involved in embryogenesis and in tumorigenesis	cSCC[2]	Ping et al., 2001 Note: in patients with history of multiple BCCS
Fas (Apo-1/CD95)	A member of the TNF-receptor superfamily. Plays important role in regulating apoptosis.	cSCC[1]	South et al., 2014
HMCN1	Encodes a large extracellular member of the immunoglobulin superfamily. Involved in anchorage of mechanosensory neurons to the epidermis, and organization of hemidesmosomes in the epidermis		South et al., 2014
FLT3	Class III receptor tyrosine kinase involved in regulation of hematopoiesis	cSCC[1]	South et al., 2014

Table 8.11 Second targeted sequencing panel gene list. Base pair mutations were seen on these genes within BCCs using whole exome sequencing (Jayaraman et al., 2014).

Gene	Gene function
TP53	Tumour suppresser gene/ apoptosis/ cell cycle arrest/ senescence
C7	A part of the complement system. Involved in formation of Membrane Attack Complex (MAC).
OR5M3	Member 3 of the olfactory receptor, family 5, subfamily M which is member of G-protein-coupled receptors (GPCR). Important in recognizing and transmission of odour signals.
UGT2B10	Involved in the elimination of potentially toxic Substances
DPP10	A member of the S9B family. Encodes a single-pass type II membrane protein.
GPR139	Belong to G protein-coupled receptors (GPCRs) family. Transmembrane signal transition.
ITIH2	A member of plasma serine protease inhibitors family. Act as stabilizer for extracellular matrix and prevent tumour invasion to the surrounding tissue
TMEM217	Information not available
PTCH1	Tumour suppressor gene. Expressed protein forms a receptor for sonic hedgehog protein which is involved in embryogenesis and in tumorigenesis
STEAP4	Present in Golgi apparatus and utilise metalloreductase function. May has role in adipocyte growth and metabolism
TMEM132D	cell-surface marker for oligodendrocyte differentiation
PPP1R3A	Regulation of carbohydrate and lipid metabolism
KCNT2	A member of potassium channel subfamily T. Associated with neurological disorders such as epilepsy.
POM121L12	Paralog to POM121C gene which is a part of nuclear pore complex (NPC).
HEATR7B2	HEAT repeat family member 7B2. SNPs on this gene are associated with Multiple sclerosis.
PCDHB3	A member of the protocadherin beta gene bundle. May be involved in cell- cell signal transduction in nerve cells.
DPCR1	Information not available
PCDHB2	A member of the protocadherin beta gene bundle. May be involved in cell- cell signal transduction in nerve cells.
NPAP1	Has biallelic expression in adult testis and brain, but paternal imprinted copy in foetal brain. Associated with Prader-Willi syndrome (imprinting disorder).
SULT1C3	A member of sulfotransferase family. It has a low affinity to alcohol associated substrates.

Table 8.12 Comparison of the percentage of gene (exons) coverage between different available whole exome capture kits. 2 lists of genes were generated according to the relevance of these genes in the pathology of NMSC and potential precancerous skin lesions (appendix table 8.10 and 8.11). **A:** The percentage of coverage to first panel gene list by three different whole exome capture kits and **B:** The percentage of coverage to the second panel gene list by the same three different whole exome capture kits (the table generated in collaboration with Dr Rueben Pengelly)**.**

A: The percentage of first panel gene list coverage by three different commercial whole exome capture kits.

Gene	Agilent SureSelect V5 % gene covered	Illumina TruSeq % gene covered	BGI % gene covered
ABCA13	100	100	100
ACACB	100	100	100
APOB	100	100	100
CACNA1B	100	100	100
CACNA1C	100	100	90.29
CDH23	100	100	95.01
CDKN2A	100	100	96.29
CNTNAP2	100	100	100
CNTNAP5	100	100	100
COL6A3	100	100	100
CSMD1	100	100	100
CSMD3	100	100	100
DMXL2	100	100	99.96
DNAH10	100	100	100
DNAH17	100	100	99.30
DNAH2	100	100	100
DNAH3	100	100	99.99
DNAH5	100	100	100
DNAH7	100	100	100
DNAH8	100	100	99.89
DST	91.91	100	100
FAT1	100	100	100
FBN1	100	100	100
FLG	100	100	100
FLNB	100	100	100
FSIP2	100	100	0
GPR158	100	100	83.52
GPR98	100	100	99.86
HMCN1	100	100	100
HRNR	100	89.82	100
KALRN	100	100	100
KIAA1211	100	100	100
LRP1B	100	100	99.88
MAP1A	100	100	100
MUC16	100	99.80	100
MUC17	100	100	100
MUC4	100	100	100
MUC5B	100	100	100
NEB	100	90.71	88.33
NOTCH1	100	100	100
NOTCH2	100	100	99.80
OBSCN	89.76	98.43	89.76
PAPPA2	100	100	99.07
PCLO	100	100	100
PKHD1	100	100	100
PRB2	100	98.64	100
RELN	100	100	100
RYR2	100	100	99.82
RYR3	100	100	100
SIGLEC1	100	100	100
SORCS3	100	100	100

A: The percentage of first panel gene list coverage by three different commercial whole exome capture kits (continue)

Gene	Agilent SureSelect V5 % gene covered	Illumina TruSeq % gene covered	BGI % gene covered
SPTA1	100	100	96.69
SRRM2	100	100	93.57
SYNE1	100	100	99.98
SYNE2	100	100	100
TEX15	100	100	100
TP53	100	100	95.41
TPO	100	100	100
UNC80	100	100	99.84
USH2A	100	100	100
XIRP2	100	100	99.80
ZFHX4	100	100	98.75
Overall average coverage	99.71	99.63	97.17

B: The percentage of second panel gene list coverage by three different commercial whole exome capture kits.

Gene	Agilent V5 % gene coverage	Illumina TruSeq % gene coverage	BGI % gene coverage
TP53	100	100	97.6
C7	100	100	100
OR5M3	100	100	99.8
UGT2B10	100	100	99.7
DPP10	100	100	99.0
GPR139	100	100	100
ITIH2	100	100	100
TMEM217	100	100	95.6
PTCH1	100	100	100
STEAP4	100	100	100
TMEM132D	100	100	100
PPP1R3A	100	100	100
KCNT2	99.8	99.8	99.8
POM121L12	100	100	100
MROH2B	100	100	100
PCDHB3	100	100	100
DPCR1	100	100	100
PCDHB2	100	100	100
NPAP1	100	100	100
SULT1C3	100	100	99.9
Overall average coverage	99.99	99.99	99.57

Table 8.13 Analysis of AK/WES data. Frequently mutated genes were arranged according to their mutation frequency in the WES results. Percentage of 100% represents genes which were mutated in all lesions and 0% is the non-mutated genes. Out of a median of 1,275 mutated genes, 94 genes were mutated in ≥ 40% of the WES data of which 43 genes are known to be mutated in cSCC. In addition, 36 genes where mutated in 20% in the study and 11 genes with no mutation in the WES have been reported to be mutated in other studies of cSCC. Frequently mutated genes were also checked if any of them had been reported in skin/ other human cancers in COSMIC database. Moreover, false positively mutated genes were picked up using the Network of Cancer Genes webpage (NGS5.0) (http://ncg.kcl.ac.uk/) (Genes were classified as false positively mutated based on the calculations in Lawrence et al. (2013)). The genes in green are the genes that were selected for target enriched sequencing.

Gene number	Gene ID	Frequency of mutated genes within the current WES AK data	Previously reported mutated gene within South et al., 2014	Previously reported mutated gene within Pickering et al., 2014	Previously reported mutated gene within Li et al., 2015	Previously reported gene within COSMIC webpage in skin cancer	Previously reported gene within COSMIC webpage other cancer	Genes reported as false positive mutated within Network of Cancer Gene
1	*TP53*	100%	64%	Yes	Yes	Yes	Yes	No
2	*TTN*	100%	≥ 40%					Yes
3	*CCDC168*	80%						Yes
4	*MUC16*	80%	≥ 40%					Yes
5	*PCLO*	80%	≥ 40%					Yes
6	*UNC79*	80%						Yes
7	*CACNA1C*	80%	≥ 40%	Yes		Yes	Yes	No
8	*ACAN*	60%				Yes	Yes	No
9	*AHNAK*	60%				Yes	Yes	No
10	*CACNA1G*	60%				Yes	Yes	No
11	*RNF103-CHMP3*	60%						No
12	*CNTNAP4*	60%				Yes	Yes	No
13	*COL24A1*	60%				Yes	Yes	No
14	*COL4A1*	60%				Yes	Yes	No
15	*CSMD1*	60%	≥ 40%			Yes	Yes	Yes
16	*CSMD3*	60%	≥ 40%			Yes	Yes	Yes
17	*CTAGE1*	60%				Yes	Yes	No
18	*DCC*	60%				Yes	Yes	No
19	*DNAH5*	40%	≥ 40%	Yes		Yes	Yes	No
20	*DNHD1*	60%				Yes	Yes	No
21	*DOT1L*	60%				Yes	Yes	No
22	*DPP10*	60%				Yes	Yes	No
23	*FBN2*	60%				Yes	Yes	No
24	*FLNC*	60%				Yes	Yes	No
25	*FREM3*	60%				No	Yes	No
26	*GPR98*	60%	≥ 40%	Yes		Yes	Yes	No
27	*HECW1*	60%				Yes	Yes	No
28	*HMCN1*	60%	≥ 40%	Yes		Yes	Yes	No
29	*HRAS*	40%	16%	Yes	Yes	Yes	Yes	No
30	*HYDIN*	60%				Yes	Yes	No
31	*KCNMA1*	60%				Yes	Yes	No
32	*KIAA0226L*	60%				Yes	Yes	No

Gene number	Gene ID	Frequency of mutated genes within the current WES AK data	Previously reported mutated gene within South et al., 2014	Previously reported mutated gene within Pickering et al., 2015	Previously reported mutated gene within Li et al., 2015	Previously reported gene within COSMIC webpage in skin cancer	Previously reported gene within COSMIC webpage other cancer	Genes reported as false positive mutated within Network of Cancer Gene
33	*LOXHD1*	60%				Yes	Yes	No
34	*LRP1B*	60%				Yes	Yes	Yes
35	*MAGEL2*	60%				Yes	Yes	Not present
36	*MGAM*	60%				Yes	Yes	No
37	*MLL2*	80%	≥ 40%	Yes	Yes	Yes	Yes	No
38	*MUC17*	60%	≥ 40%			Yes	Yes	Yes
39	*MUC2*	60%					Yes	No
40	*MYH1*	60%				Yes	Yes	No
41	*MYO16*	60%				Yes	Yes	No
42	*NOTCH2*	60%	63%	Yes	Yes	Yes	Yes	No
43	*OR13C5*	60%				Yes	Yes	No
44	*OR4N4*	60%				Yes	Yes	No
45	*PAPPA2*	60%	≥ 40%	Yes		Yes	Yes	No
46	*PDZD2*	60%				Yes	Yes	No
47	*PKHD1L1*	60%				Yes	Yes	No
48	*PLXNA2*	60%				Yes	Yes	No
49	*POLR1A*	60%				Yes	Yes	No
50	*PRUNE2*	60%				Yes	Yes	No
51	*RIMS1*	60%				Yes	Yes	No
52	*RP1*	60%				Yes	Yes	No
53	*RYR2*	60%	≥ 40%			Yes	Yes	Yes
54	*RYR3*	60%	≥ 40%			Yes	Yes	No
55	*SCN1A*	60%				Yes	Yes	No
56	*SGK223*	60%				Yes	Yes	No
57	*SLC44A5*	60%				Yes	Yes	No
58	*SPECC1*	60%					Yes	No
59	*SPEG*	60%				Yes	Yes	No
60	*SPTB*	60%				Yes	Yes	Yes
61	*SYT4*	60%				Yes	Yes	No
62	*SZT2*	60%				Yes	Yes	No
63	*TEX15*	60%	≥ 40%	Yes		Yes	Yes	No
64	*TMEM132C*	60%				Yes	Yes	No
65	*TRRAP*	60%				Yes	Yes	No
66	*ZNF521*	60%				Yes	Yes	No
67	*ZNF646*	60%				Yes	Yes	No
68	*ZNF775*	60%				Yes	Yes	No
69	*FAT1*	60%	≥ 40%	Yes		Yes		No
70	*APOB*	40%	≥ 40%	Yes		Yes	Yes	No

Gene number	Gene ID	Frequency of mutated genes within the current WES AK data	Previously reported mutated gene within South et al., 2014	Previously reported mutated gene within Pickering et al., 2015	Previously reported mutated gene within Li et al., 2015	Previously reported gene within COSMIC webpage in skin cancer	Previously reported gene within COSMIC webpage other cancer	Genes reported as false positive mutated within Network of Cancer Gene
71	*CNTNAP5*	40%	≥ 40%			Yes	Yes	No
72	*DNAH10*	40%	≥ 40%			Yes	Yes	No
73	*DNAH17*	40%	≥ 40%			Yes	Yes	No
74	*DNAH2*	40%	≥ 40%			Yes	Yes	No
75	*DNAH3*	40%	≥ 40%			Yes	Yes	No
76	*DNAH7*	40%	≥ 40%			Yes	Yes	No
77	*DNAH8*	40%	≥ 40%			Yes	Yes	No
78	*ABCA13*	40%	≥ 40%			Yes	Yes	No
79	*FBN1*	40%	≥ 40%			Yes	Yes	No
80	*FLT3*	40%	≥ 40%	Yes		Yes	Yes	No
81	*GPR158*	40%	≥ 40%			Yes	Yes	No
82	*MAP1A*	40%	≥ 40%			Yes	Yes	No
83	*MLL3*	40%	≥ 40%			Yes	Yes	No
84	*NOTCH1*	40%	75%	Yes	Yes	Yes	Yes	No
85	*PKHD1*	40%	≥ 40%			Yes	Yes	No
86	*SORCS3*	40%	≥ 40%			Yes		Not present
87	*SPTA1*	40%	≥ 40%			Yes	Yes	No
88	*SRRM2*	40%	≥ 40%			Yes	Yes	Not present
89	*TET2*	40%				Yes	Yes	No
90	*USH2A*	40%	≥ 40%			Yes	Yes	Not present
91	*XIRP2*	40%	≥ 40%			Yes	Yes	No
92	*ZFHX4*	40%	≥ 40%			Yes	Yes	No
93	*ACACB*	20%	≥ 40%			Yes	Yes	No
94	*AJUBA*	20%		Yes		Yes	Yes	Not present
95	*ARID5B*	20%			Yes	Yes	Yes	Not present
96	*BBS9*	20%		Yes		Yes	Yes	Not present
97	*C1orf173*	20%	≥ 40%			Yes	Yes	No
98	*CACNA1B*	20%	≥ 40%			Yes	Yes	No
99	*CARD11*	20%		Yes		Yes	Yes	No
100	*CDH23*	20%	≥ 40%			Yes	Yes	No
101	*CDKN2A*	20%	23%	Yes	Yes	Yes	Yes	No
102	*COBLL1*	20%		Yes		Yes	Yes	Not present
103	*COL6A3*	20%	≥ 40%			Yes	Yes	No
104	*CREBBP*	20%			Yes	Yes	Yes	No
105	*DCLK1*	20%		Yes		Yes	Yes	Not present

Gene number	Gene ID	Frequency of mutated genes within the current WES AK data	Previously reported mutated gene within South et al., 2014	Previously reported mutated gene within Pickering et al., 2015	Previously reported mutated gene within Li et al., 2015	Previously reported gene within COSMIC webpage in skin cancer	Previously reported gene within COSMIC webpage other cancer	Genes reported as false positive mutated within Network of Cancer Gene
106	*DCLRE1A*	20%		Yes		Yes	Yes	Not present
107	*DST*	20%	≥ 40%			Yes	Yes	No
108	*EP300*	40%	≥ 40%		Yes	Yes	Yes	No
109	*FLG*	20%	≥ 40%	Yes		Yes	Yes	No
110	*KIAA1211*	20%	≥ 40%			Yes	Yes	No
111	*KRAS*	20%	13%	Yes		Yes	Yes	No
112	*MTOR*	20%			Yes	Yes	Yes	Not present
113	*MUC4*	20%	≥ 40%			Yes	Yes	No
114	*MUC5B*	20%	≥ 40%			Yes	Yes	No
115	*NEB*	20%	≥ 40%			Yes	Yes	No
116	*NRAS*	20%	5%	Yes		Yes	Yes	No
117	*OBSCN*	20%	≥ 40%			Yes	Yes	No
118	*PARD3*	20%		Yes		Yes	Yes	Not present
119	*PRB2*	20%	≥ 40%			Yes	Yes	No
120	*PTCH1*	40%	≤40%	Yes		Yes	Yes	No
121	*RBM46*	20%		Yes		Yes	Yes	Not present
122	*RELN*	20%	≥ 40%			Yes	Yes	No
123	*RIPK2*	20%		Yes		Yes	Yes	Not present
124	*RIPK4*	20%			Yes	Yes	Yes	Not present
125	*SMARCA4*	20%		Yes		Yes	Yes	No
126	*SYNE1*	20%	≥ 40%			Yes	Yes	No
127	*SYNE2*	20%	≥ 40%			Yes	Yes	No
128	*TERT*	20%		Yes		Yes	Yes	No
129	*ZNF544*	20%		Yes		Yes	Yes	No
130	*UNC80*	20%	≥ 40%			Yes	Yes	No
131	*CNTNAP2*	0	≥ 40%			Yes	Yes	Yes
132	*DMXL2*	0	≥ 40%			Yes	Yes	Not present
133	*EGFR*	0		Yes		Yes	Yes	No
134	*EZH2*	0		Yes		Yes	Yes	No
135	*FSIP2*	0	≥ 40%			Yes	Yes	Yes
136	*HRNR*	0	≥ 40%			Yes	Yes	Yes
137	*KNSTRN*	0				Yes	Yes	Not present
138	*PCDHA4*	0	≥ 40%			Yes	Yes	Yes
139	*RYR1*	0	≥ 40%			Yes	Yes	No
140	*SIGLEC1*	0	≥ 40%			Yes	Yes	No
141	*WT1*	0		Yes		Yes	Yes	No

Table 8.14 WES AK data on the 18 genes selected for target enriched sequencing. The number and the position of the amino acid changes on resultant protein from each of the 18 genes from figure 3.28 are shown in this table. M - missense changes, N - nonsense changes and F - frameshift mutations.

Gene ID	patient ID				
	28	30	51	54	68
CACNA1C	P1785S:M P1779S:M P1827S:M P1798S:M	P1807S:M P1820S:M P1787S:M P1776S:M P1768S:M	W1368R:M W1399R:M W1351R:M W1379R:M W1373R:M W1371R:M W1340R:M	G250R:M	
CDKN2A	R129C:M P114L:M				
DNAH5		D3236N:M D2817N:M G2684E:M D3236N:M	S987F:M S1269F:M V3245M:M E2814K:M		
EP300	P666L:M P666S:M		P666S:M		
FAT1		D1484N:M F614L:M V482I:M S404R:M	S1310L:M N1556K:M V482I:M S404R:M		S3554A:M R1064G:M Q2933P:M V862L:M
FLT3	G139R:M			G139R:M V641I:M	
GPR98	M1263I:M M1262*:N M1265*:N	P1898S:M	G533S:M E768K:M G533D:M D582G:M		
HMCN1	G1852E:M	Q1982*:N S842N:M	E3642K:M		
HRAS	G13D:M G162*:N	G168R:M Q61L:M			
KRAS		G12D:M			
MLL2	Q3788*:N Q1522*:N R5448*:N	S4560F:M	D5044N:M		S4560F:M
NOTCH1	Q513*:N E450K:M	E450K:M			
NOTCH2	E1526K:M E1520*:N		D2037N:M P2219L:M		E2258K:M M2257l:M M2254fs:F
NRAS		M111T:M			
PAPPA2	G1752E:M		N240I:M	D921N:M	
PTCH1	G15D: M G15S: M G100D: M G100S: M	G165D: M G165S: M G166D: M G166S: M			
TEX15	D2585N:M		S1352F:M	S195R:M	
TP53	R282fs:F R150fs:F R243fs:F	S183*:N S51*:N S144*:N	P322fs:F P190fs:F P283fs:F	R342*:N R303*:N R210*:N	K141Q:M K136Q:M Y236D:M Y104D:M Y197D:M

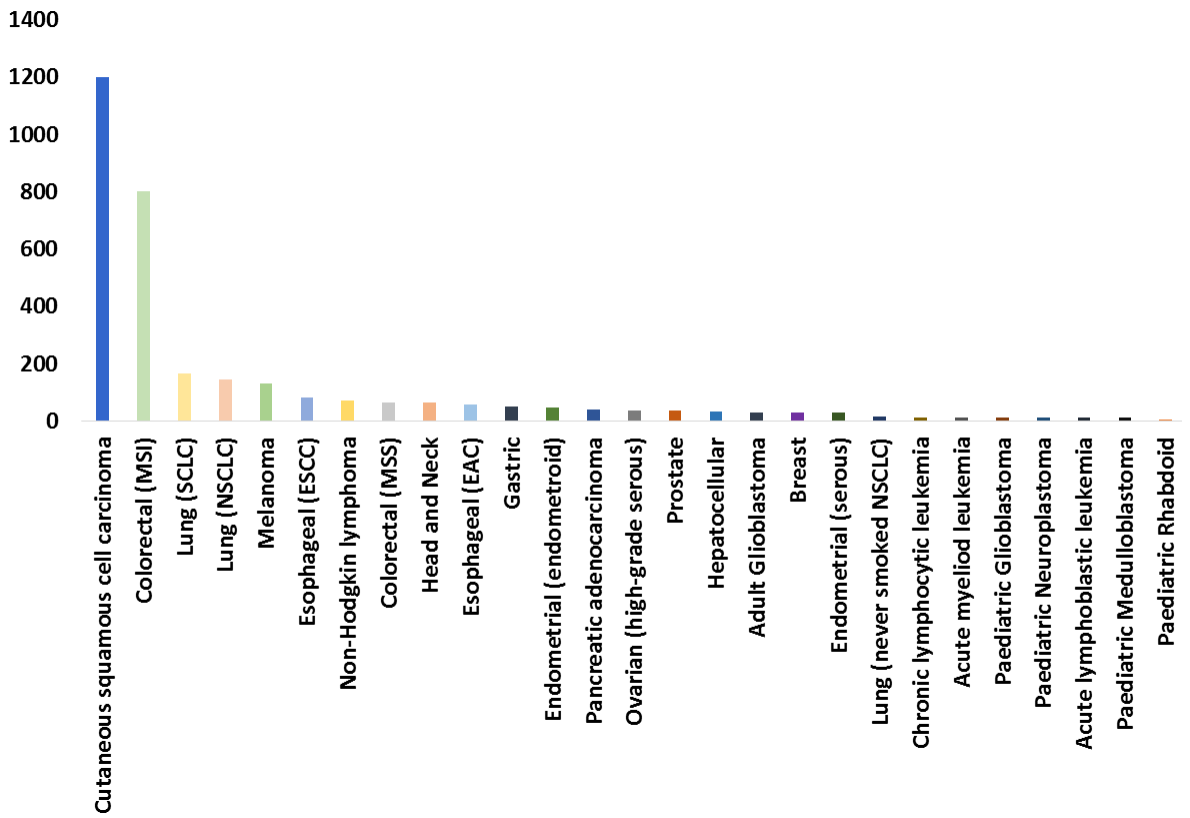


Figure 8.1 Comparison of mutation burden between different malignancies. The median number of nonsynonymous mutations per malignancy (using WES) in different common malignancies. X-axis ordered by median number of mutations across different lesions (left to right). Adapted from South et al., 2014 based on information from Vogelstein et al., 2013.

Table 8.15 Selected samples for target enriched sequencing (sAK, sBD, adjacent AK/cSCC and adjacent BD/cSCC) lesions. Patient's ID represents the order of the selected samples from the database search. Abbreviations used include M = Male, F = Female. sAK = solitary AK. sBD = solitary Bowen's disease. cSCC = cutaneous squamous cell carcinoma.

	Patient ID	Year of birth	Age at biopsy	Gender	Site of the lesion	Year of biopsy
sAK	2	1915	88	M	Forehead	2003
	3	1943	61	M	Forehead	2004
	5	1932	71	M	Left ear	2003
	6	1912	91	M	Scalp	2003
	13	1951	54	M	Neck	2005
	15	1930	76	M	Shoulder/chest	2006
	16	1922	84	M	Scalp	2006
	20	1941	65	F	Scalp	2006
	25	1941	72	M	Right eyebrow	2013
	27	1933	80	M	Left ear	2013
	29	1932	81	M	Right scalp	2013
	31	1940	73	M	Left ear	2013
	37	1936	77	M	Left pre-auricular area	2013
	42	1934	79	M	Dorsum of the left hand.	2013
	45	1921	92	F	Right arm	2013
	52	1935	78	M	Left cheek	2013
	53	1945	68	M	Neck	2013
	56	1941	72	M	Right middle finger	2013
	62	1948	65	F	Nape of the neck	2013
	64	1922	91	M	Scalp	2013
	69	1934	79	M	Left elbow	2013
	71	1943	70	F	Left lower lid	2013
	72	1939	81	M	Left inner knee	2012
	73	1937	78	F	Upper left back	2013
	74	1947	65	F	Right middle finger	2013
sBD	1	1921	93	M	Scalp	2014
	2	1928	86	F	Neck	2014
	3	1920	94	F	Scalp	2014
	4	1945	69	M	Left hand	2014
	5	1936	78	F	Right lower leg	2014
	6	1939	75	F	Left lower leg	2014
	7	1943	71	M	Right ear	2014
	8	1948	66	M	Right Cheek	2014
	9	1928	86	M	Scalp	2014
	10	1932	82	F	Left hand	2014
	11	1941	73	M	Scalp	2014
	12	1929	85	F	Scalp	2012
	13	1957	57	F	Left Cheek	2014
	14	1930	84	M	Left check	2014
	15	1946	68	M	Left check	2014
	16	1930	84	F	Right hand	2014
	17	1929	85	M	Scalp	2014
	18	1928	86	M	Left ear	2014
	19	1929	85	F	Right lower leg	2014
	20	1930	84	M	Left ear	2014
	21	1927	87	M	Left Cheek	2014
	22	1930	84	M	Left Cheek	2013
	23	1934	80	M	Scalp	2014
	24	1926	88	M	Scalp	2013
	25	1936	78	M	Right hand	2014
	26	1929	85	F	Left lower leg	2014
	28	1931	83	F	Right lower leg	2014
	29	1927	87	F	Neck	2014
	30	1944	70	M	Scalp	2014

	Patient ID	Year of birth	Age at biopsy	Gender	Site of the lesion	Year of biopsy
AK/cSCC	3	1927	84	M	Scalp	2011
	4	1948	65	M	Scalp	2013
	5	1940	74	M	Left hand	2014
	6	1928	85	M	Scalp	2013
	7	1937	77	M	Left cheek	2014
	8	1933	81	M	Right elbow	2014
	10	1925	88	F	Nasal bridge	2013
BD/cSCC	11	1928	86	M	Scalp	2014
	12	1936	78	M	Scalp	2014
	13	1925	89	M	Nose	2014
	14	1946	68	M	Scalp	2014
	15	1944	70	M	Nose	2014
	16	1945	69	M	Scalp	2014
	17	1957	57	M	Right cheek	2014
	18	1920	94	M	Right cheek	2014

Table 8.16 The total amount of DNA obtained from lesions selected for target enriched sequencing. DNA concentration was measured using Qubit 2.0 fluorometer and expressed in ng. AK = Actinic keratosis. BD = Bowens disease. cSCC = Cutaneous squamous cell carcinoma. NS = Normal skin. ng = nanogram.

Sample ID	Total amount of DNA within 15 µl volume expressed in ng									
	Solitary AK		Solitary BD		Adjacent AK/cSCC			Adjacent BD/cSCC		
	Lesion	NS	Lesion	NS	cSCC	AK	NS	cSCC	BD	NS
1			353.1	645.1						
2	215	119	475.7	674.7						
3	435	158.75	432.6	622.1	480	505	670			
4			497.5	631	717.5	307.5	214.25			
5	435	400	468.1	482.9	497.5	415	915			
6	295	287.5	455.964	186.124	2467.5	1897.5	812.5			
7			416.558	502.26	171	152.25	270			
8			377.928	366.217	2237.5	257.5	1075			
9			409.732	428.15						
10			444.535	458.501	1925	1208.5	745			
11			423.618	546.9				925	975	400
12			463.338	302.816				995	822.5	557.5
13	549	512.5	405.84	496.484				230.25	212.5	297.5
14			503.276	486.389				515	300	155.5
15	319	725	390.762	415.641				337.5	482.5	150
16	529	405	514.593	847.121				845	507.5	475
17			446.807	679.318				580	452.5	320
18			424.653	341.475				837.5	272.5	332.5
19			423.401	729.04						
20	567.5	442.5	540.549	420.465						
21			467.351	305.31						
22			366.032	354.909						
23			370.066	553.335						
24			362.007	488.982						
25	330	882.5	297.907	760.626						
26			512.389	310.059						
27	375	375	434.763	232.265						
28			392.426	245.766						
29	385	980	344.249	128.336						
30			353.1	645.1						
31	201.5	672.5								
37	486	812.5								
42	326	300								
45	355	297.5								
52	385	210.25								
53	280	445								
56	592.5	400								
62	223.25	445								
64	260	810								

Sample ID	Total amount of DNA within 15 µl volume expressed in ng									
	Solitary AK		Solitary BD		Adjacent AK/cSCC			Adjacent BD/cSCC		
	Lesion	NS	Lesion	NS	cSCC	AK	NS	cSCC	BD	NS
69	300	450								
71	476	420								
72	419	927.5								
73	345	233.25								
74	548	622.5								

Table 8.17 Characterisation of subjects stained positive for PIP and selected for target enriched sequencing. Patient's ID represents the order of the samples as were received. The age, gender and site of the lesion were included. Abbreviations M=Male, F=Female.

Patient ID	Number of selected PIPs per subject	Selected PIP ID	Date of birth	Age at biopsy	Gender	Site of the lesion	Year of biopsy
13	1	1	1927	87	M	Right hand	2014
28	4	1	1937	77	F	Left hand	2014
		2					
		3					
		4					
66	1	1	1934	81	M	Left side of scalp	2015
77	1	1	1932	83	F	Upper chest	2015
87	3	1	1948	67	F	Left cheek	2015
		2					
		3					
93	1	1	1932	83	M	Right temple	2015
98	1	1	1936	79	M	Right shoulder	2015
102	1	1	1927	88	M	Shin of the leg	2015
103	1	1	1937	78	F	Shin of the leg	2015
104	1	1	1929	86	F	Left cheek	2015
107	1	1	1933	82	F	Right forearm	2015
110	1	1	1921	94	M	Abdomen	2015
111	1	1	1932	83	F	Left temple	2015
113	1	1	1942	73	F	Left forearm	2015
126	4	1	1947	68	M	Left helix	2015
		2					
		3					
		4					

Table 8.18 DNA concentration of PIP samples selected for target enriched sequencing. PIPs were identified and dissected. Then DNA was extracted from both the dissected PIP and p53 negative epidermis of the same patient. Extracted DNA concentration was measured using Qubit 2.0 fluorometer and expressed in both ng/µl and as a total amount in ng within total volume of 15 µl.

Sample ID	DNA concentration			
	p53 negative normal skin		p53 immunopositive patch (PIPs)	
	ng/µl	Total amount in ng	ng/µl	Total amount in ng
PIP-13	14.2	213	12.00	180
PIP-28	46.5	697.5		
PIP-28-1			6.8	102
PIP-28-2			7.1	106.5
PIP-28-3			16.3	244.5
PIP-28-4			25	375
PIP-66	46.9	703.5	18.6	279
PIP-77	29.4	441	10.1	151.5
PIP-87	51.4	771		
PIP-87-1			8.89	133.3
PIP-87-2			9.10	136.5
PIP-87-3			16.5	247.5
PIP-93	39.6	594	10.3	154.5
PIP-98	31.1	466.5	19.8	297
PIP-102	34.9	523.5	16.1	241.5
PIP-103	36.7	550.5	15.7	235.5
PIP-104	58.0	870	31.3	469.5
PIP-107	30.0	450	36.9	553.5
PIP-110	55.06	825.9	23.4	351
PIP-111	14.9	223.5	37.3	559.5
PIP-113	27.7	415.5	7.43	111.4
PIP-126	46	690		
PIP-126-1			29	435
PIP-126-2			20	300
PIP-126-3			11	165
PIP-126-4			37	555