Blood First Edition Paper, prepublished online May 18, 2018; DOI 10.1182/blood-2018-01-829291

TO THE EDITOR:

Single cell exomes in an index case of amp1q21 multiple myeloma reveals more diverse mutanomes than the whole population

Dean Bryant,[1] Will Tapper,[2] Nicola J Weston-Bell,[1] Arnold Bolomsky,[3] Li Song,[4] Shengtao Xu,[4] Andrew R Collins,[2] Niklas Zojer[3] & Surinder Singh Sahota[1]

[1]Tumour Immunogenetics Group, Cancer Sciences Academic Unit, Faculty of Medicine, University of Southampton, United Kingdom; [2]Human Health and Development Academic Unit, Faculty of Medicine, University of Southampton, United Kingdom; [3]Department of Medicine I, Wilhelminenspital, Wilhelminen Cancer Research Institute, Vienna, Austria; [4]BGI Genomics, BGI-Shenzhen, Shenzhen 518083, China

A focus on defining genomic mutations at the whole population level in cancer may underestimate clonal plasticity in the tumor genome, and how it generates genetic variation between individual single cells (SCs). This may impact on understanding cancer evolution and biological behaviour. Differences between SCs are known to underpin intraclonal variation (ICV), considered a major driver in cancer progression where clonal competitor cells evolve by Darwinian mechanisms.[1] Knowledge of the entire spectrum of genome-wide mutations in SCs however, requires an unbiased and unselected approach. In this, SC whole exome sequencing (WES) appears essential to understand functional gene variation in the tumour clone. To probe this, we examined SC WES in an index case of amp1q21 multiple myeloma (MM).

MM is a lethal plasma cell malignancy locating in the bone marrow. Significant efforts have now determined the nature of genome-wide somatic mutations in MM, but this work has largely focused at the whole tumor population level. At this level, of the top 20 mutated genes the most frequently mutated are *NRAS* (19%) and *KRAS* (16%),[2] and no universally shared gene mutations have emerged in MM, revealing a marked disease heterogeneity. The extent of ICV in these genomic data in individual tumors has been inferred by mapping subclonal mutations using variant allele frequencies (VAFs), defined as the percentage of total reads derived from the variant allele, obtained from genome sequencing of a single pool of thousands to millions of purified tumor cells.[3] VAFs in bulk tumor population genomic sequencing data however do not allow identification of SC mutations. The first comprehensive profile of mutations in MM SCs employed a pre-selected targeted gene sequencing approach, in which 35 variant loci of 13 commonly mutated genes in MM tumor bulk populations were assessed in SC amplified DNA using massively parallel genome sequencing.[4] This notable study reproducibly captured the 13 gene mutations with a detection sensitivity of ~93% in

MM cell line derived SCs, but was essentially aimed at evaluating the utility of targeted genome sequencing of individual circulating tumor cells to assess disease status.[4] Here we report unbiased WES in SCs from an index case of amp1q21 MM at disease presentation.

We performed WES in MM bulk population (1000 cells) and SCs (n=20) levels respectively, comparing data with autologous germline T-cell genomes (bulk 1000 cells and n=5 SCs). Full details are provided in Data Supplement (DS). Bulk 1000 cells and SCs were isolated from CD138+CD38++ or CD3+ FACS sorted tumor population and T-cells separately by precision micropipetting, and isolated cell DNA subjected to whole genome amplification (WGA), library preparation and Illumina sequencing Genome sequencing data underwent concise bioinformatics analyses and data was determined as high quality (Figure S1; Table 1).

Sequencing was performed to a mean depth of 57.7x, achieving a mean of 82.1% coverage of target exome at ≥5x, mean allele dropout (ADO) rate of 31.27% and a mean het/hom ratio of 0.72 (DS; Table S1). These data surpassed the SC WES mean read depths of ~30-33x and ~70-80% exome coverage at ≥ 5x reported in early seminal SC WES studies.[5,6]

We next assessed chromosomal amplifications and deletions from WES data by low resolution copy number variant (CNV) analysis (Figure S2). There was no apparent evidence of subclonal variation in the karyotype of tumor cells, suggesting genomic aberrations as early events in tumorigenesis in MM. Loss of heterozygosity (LOH) were not called as they were indistinguishable from ADO. These observations indicate that karyotype determined by CNV does not segregate ICV at the SC level from WES data.

Somatic variants were then identified in WES data by pairwise comparison of bulk tumor population with bulk germline T-cells, excluding those present in databases of human germline variation. Stringent quality control measures were employed to reduce the number of false positive calls (DS), and we identified 48 high confidence variants in the bulk tumor population that were considered as acquired somatic variants (Table 1; Table S1). Of these 48 variants in the tumor bulk exome, only 34 (71%) were identified in varying penetration as mutations in SCs, commensurate with VAFs that indicated ICV (Figure 1A,B). We also determined variation between SCs in principal component analysis plots (PCA; data not shown), providing further evidence of intraclonal heterogeneity (Table S2). Overall however, 29% of bulk population somatic variants were not seen in SCs. A panel of mutated genes (labelled P1 on Figure 1B, spanning genes *RNF112* to *CBX6*) present in the bulk

population were notably absent in several SCs despite VAFs >40% (Table S1), and draw caution in inferences from gene VAFs in bulk tumor population genomic data.

We also identified an additional 21 somatic variants that were only called in 2 or more SC exomes but not in the bulk tumor population despite high coverage at each of 17 sites in bulk exome data (mean sequence depth 105x; range 36-330x). The vast majority of the 21 variants thereby provided robust evidence of somatic mutations occurring in differing SCs that were not detectable at the whole tumor population level, and revealing a greater sensitivity in identifying mutations when using SC genome-wide methodologies. The somatic variants identified in SC exomes confirmed that mapping subclonal mutations across a MM tumor population requires an unbiased approach.

To confirm variant calls and occurrence of mutations in our study, 15 randomly selected nonsynonymous SNVs were validated by Sanger sequencing in the bulk tumor population and 4 randomly chosen SCs. 15/15 variants were confirmed in the bulk tumor amplified DNA and in 55/55 variants amplifiable in SC DNA, to obtain 100% concordance (Table S3). An additional 10 nonsynonymous SNVs that were only observed in SCs were also re-sequenced, and 7/10 variants were amplified in at least one SCs, and in each the variant nucleotide confirmed by sequence. By Sanger sequencing, we confirmed somatic variants with VAF ranging from 19-100%, of which 7/15 had a VAF <40%.

We identified a total of 69 unique mutations present with high confidence in the bulk and/or SC tumour exomes, and 48 of these were predicted as potentially deleterious (nonsynonymous, stopgain, frameshift indel) and 37/48 were predicted to be of functional relevance to disease origins (Table S2). All but 2 deleterious variants occurred in genes also reported in COSMIC.[1] Driver gene analysis using IntOgen and KEGG pathway analysis identified five potential driver genes (*ANK3*, *AXIN1, BRCA2, MAP4K3*, *TRIP10*), but these appeared subclonal (Figure 1B).

Strikingly, these SC WES data indicate that the mutational status of the MM genome is markedly underestimated by bulk tumor population analysis, and overlooked by targeted gene resequencing approaches. The 21 somatic variants identified in 2 or more SC exomes reveal replicates and an apparent ~44% increase in mutational complexity not suspected from 48 bulk population mutations. Importantly, our data in a lymphoid malignancy are comparable to findings from 16 SC nucleus-based exomes in a triple-negative breast cancer case, where in 2 or more SC exomes, 145 non-

synonymous mutations were identified that were not detected at bulk population level (374 variants), an apparent ~40% increase.[7]

The implications of our data in MM follow. We compared 20 SCs with bulk tumor (1000 cells), evaluating only 20/1000 (0.02%) cells, yet the magnitude of how mutational load is underestimated emerges. The scale of the problem and the need to derive a truly representational map of SC exomes is revealed by considering estimates of tumor mass in MM. It has been shown using ^{125}I loaded synthesis of IgG and mathematical modelling that MM tumor mass approaches 10^{12} cells.[8,9] Our case was an IgG MM patient with 80-90% bone marrow disease infiltrate at presentation, with potentially a comparable tumor mass. It becomes a major challenge then to map the functional SC genome in a tumor mass of this scale. This is an essential requirement to fully understand how cell-to-cell variation determines tumor biology and behaviour, and to develop individualised medicine. Any non-synonymous somatic mutation in a SC has considerable scope to alter cellular function, intrinsically and extrinsically by altering molecular cross-talk with microenvironment cells. The sum of variation in SC behaviour will determine population level growth and progression. Singular findings from SC exomes also challenge the concept of the tumor clone, with a key question being how many SCs are identical and how this will relate to defining a clone.[7] Although a study of an index case of 1q21 MM, we have nevertheless begun to unravel the true genomic complexity of this lethal disease at the SC level.

Acknowledgements

Funded by Bloodwise (United Kingdom). The authors acknowledge use of the IRIDIS High Performance Computing Facility and associated support services at the University of Southampton in the completion of this work.

Authorship

DB: performed research, collected data, analysed and interpreted data, performed bioinformatics and statistical analysis, contributed to writing the manuscript. WT, ARC: performed bioinformatics and statistical analysis, analysed data. NWB, AB, NZ: sample preparations, analysed and interpreted data (NZ also arranged Patient Consent and Ethics approval in Vienna, Austria; contributed to design of study). LS, SX: co-ordinated WES and data preparation by BGI. SSS: designed research, analysed and interpreted data, wrote the manuscript. All authors approved the final submission of manuscript.

Conflict-of-Interest disclosure

All authors report no conflict of interest in relation to this work.

Corresponding author:

Dr Surinder Singh Sahota, Somers Cancer Research Building (MP891), Southampton University Hospitals, Tremona Road, Southampton SO16 6YD, UK; E-mail: s.s.sahota@soton.ac.uk ; Phone: +44 (0)23 8120 8768; Fax: +44 (0)23 8120 5152

Footnote

This article contains a data supplement.

REFERENCES

1. McGranahan N, Swanton C. Biological and Therapeutic Impact of Intratumor Heterogeneity in Cancer Evolution. Cancer Cell. 2015;27(1):15-26.

2. COSMIC: Catalogue of Somatic Mutations in Cancer. 2017. *Cosmic Cancer Browser*. [ONLINE] Available at: http://cancer.sanger.ac.uk/cosmic/browse/tissue#in=t&sn=haematopoietic_and_lymphoid_tissue&ss=NS&hn=lymphoid_neoplasm&sh=plasma_cell_myeloma. [Accessed 31 August 2017].

3. Bolli N, Avet-Loiseau H, Wedge DC, Van Loo P, Alexandrov LB, Martincorena I, et al. Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. Nature Communications. 2014;5:2997.

4. Lohr JG, Kim S, Gould J, Knoechel B, Drier Y, Cotton MJ, et al. Genetic interrogation of circulating multiple myeloma cells at single-cell resolution. Science Translational Medicine. 2016;8(363):363ra147-363ra147.

5. Hou Y, Song L, Zhu P, Zhang B, Tao Y, Xu X, et al. Single-cell exome sequencing and monoclonal evolution of a JAK2-negative myeloproliferative neoplasm. Cell. 2012;148(5):873-85.

6. Xu X, Hou Y, Yin X, Bao L, Tang A, Song L, et al. Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. Cell. 2012;148(5):886-95.

7. Wang Y, Waters J, Leung ML, Unruh A, Roh W, Shi X, et al. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. Nature. 2014;512(7513):155-60.

8. Salmon SE, Smith BA. Immunoglobulin synthesis and total body tumor cell number in IgG multiple myeloma. J Clin Invest. 1970;49:1114-1121.

9. Sullivan PW, Salmon SE. Kinetics of tumor growth and regression in IgG multiple myeloma. J Clin Invest. 1972;51:1697-1708.

Table 1: Variant counts identified in tumor bulk population and single cell exomes

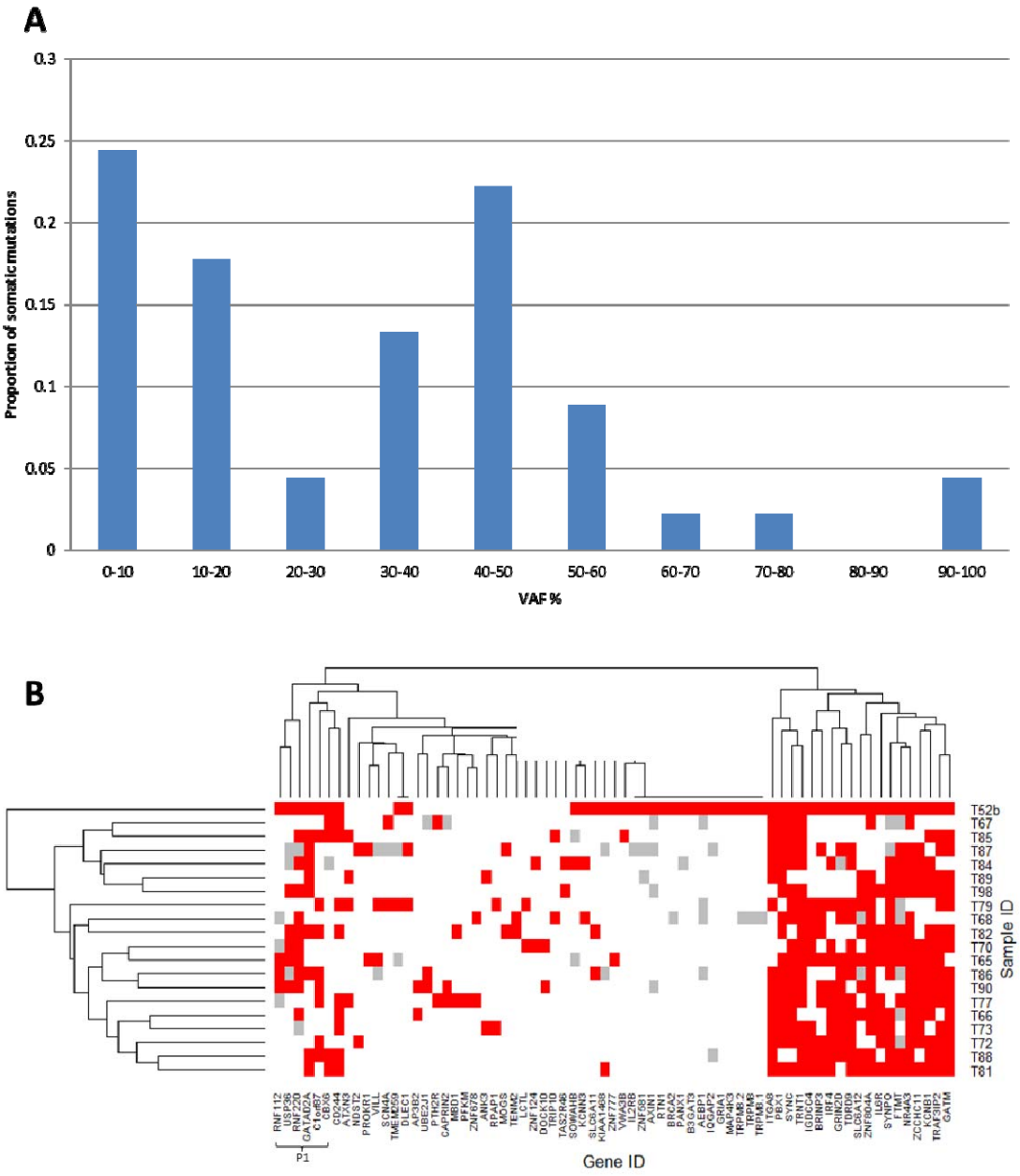
Variants called*	Nonsynonymous SNV	Synonymous SNV	Stopgain SNV	Frameshift insertion	Frameshift deletion	Total
Bulk exome	20	14	1	9	3	48
• *In 1+ SC*	*16*	*10*	*1*	*6*	*1*	*34*
• *Not in 1+ SC*	*4*	*4*	*0*	*4*	*2*	*14*
Total 2+ single cells	28	16	3	3	1	51
• *In bulk*	*16*	*9*	*1*	*3*	*1*	*30*
• *Not in bulk*	*12*	*7*	*2*	*0*	*0*	*21*
Total unique	32	21	3	10	3	69

*Variant counts of each type of are tabulated for the bulk exome and SC exomes. For the SC exomes, selection criteria required that the variant be present in 2+ SC cells to be counted. The number of variants identified in 2+ SC exomes that were also called (n=30) or not called (n=21) in the bulk exome are also given.

FIGURE LEGEND

Figure 1. Profiles of somatic variants in the bulk tumor population and SC exomes in an index case of amp1q21 MM. **Panel A**: Variant allele frequency (VAF) of 46 high confidence somatic variants in the bulk tumor population exome. **Panel B**: Two-way unsupervised clustering using Euclidean distance and complete linkage. Sample T52b (top) is MM bulk population, and aligned below are representation of SC exome data derived from individual tumor cells. The genotypes (homozygous reference = 0, white [germline], heterozygous/homozygous alternative = 1, red [mutation]) of 69 mutations were used as input (high confidence somatic from the bulk and 2+SC exomes). Acquired somatic mutations in tumor cells (red) reveal the span of mutational complexity in SCs not apparent in bulk tumor population. 49 genotypes were missing due to lack of coverage (grey). Genes *RNF112* to *CBX6,* labelled P1 (bottom left) are examples of variants called in the bulk tumor exome but notably absent in many SCs.

Figure 1



Prepublished online May 18, 2018;
doi:10.1182/blood-2018-01-829291

Single cell exomes in an index case of amp1q21 multiple myeloma reveals more diverse mutanomes than the whole population

Dean Bryant, Will Tapper, Nicola J. Weston-Bell, Arnold Bolomsky, Li Song, Shengtao Xu, Andrew R. Collins, Niklas Zojer and Surinder Singh Sahota

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036.