Recurrent activating STAT5B N642H mutation in myeloid neoplasms
with eosinophilia

Nicholas C. P. Cross,1,2 Yvette Hoade,1,2 William J. Tapper,1 Gonzalo Carreno-Tarragona,1,2 Tiziana Fanelli,3 Mohamad Jawhar,4 Nicole Naumann,4 Iwo Pieniak,1 Johannes Lübke,4 Sahra Ali,5 Kaljit Bhuller,6 Sonja Burgstaller,7 Catherine Cargo,8 Jamie Cavenagh,9 Andrew S. Duncombe,10 Emma Das-Gupta,11 Paul Evans,8 Peter Forsyth,12 Philip George,11 Charlotte Grimley,11 Fergus Jack,13 Laura Munro,14 Varun Mehra,15 Kavita Patel,16 Ali Rismani,17 Gabriela Sciuccati,18 Rowena Thomas-Dewing,19 Patrick Thornton,20 Andres Virchis,21 Simon Watt,22 Louise Wallis,23 Alastair Whiteway,24 Kris Zegocki,25 Barbara J. Bain,26 Andreas Reiter4 & Andrew Chase1,2

1 Faculty of Medicine, University of Southampton, Southampton, UK
2 Wessex Regional Genetics Laboratory, Salisbury NHS Foundation Trust, Salisbury, UK
3 Center Research and Innovation of Myeloproliferative Neoplasms, AOU Careggi, University of Florence, Firenze, Italy
4 University Hospital Mannheim, Heidelberg University, Mannheim, Germany
5 Hull & East Yorkshire Hospitals NHS Trust, Hull, UK
6 University Hospitals of Leicester NHS Trust, Leicester, UK
7 Klinikum Wels-Grieskirchen, Wels, Austria
8 HMDS, St. James’s University Hospital, Leeds, UK
9 St Bartholomew’s Hospital, London, UK
10 University Hospitals Southampton, Southampton, UK
11 Nottingham University Hospitals NHS Trust, Nottingham, UK
12 Raigmore Hospital, Inverness, UK
13 Poole Hospital NHS Trust, Poole, UK
14 York Teaching Hospital NHS Trust, York, UK
15 King’s College Hospital, London, UK
16 Mid Yorkshire Hospitals NHS Trust, UK
17 Whittington Health & University College London Hospitals, London, UK
18 Hospital de Pediatria "Prof. Dr. Garrahan", Buenos Aires, Argentina
19 Salford Royal Hospital, Salford, UK
20 Beaumont and Connolly Hospitals, Dublin, Ireland
21 Barnet & Chase Farm Hospitals, Wellhouse Lane, Barnet, UK
22 Manchester University NHS FT, Manchester, UK
23 Royal Bournemouth Hospital, Bournemouth, UK
24 Southmead Hospital, Bristol, UK
25 Whipps Cross University Hospital, London, UK
26 St Mary’s Hospital, London, UK.
Correspondence to:

Professor N.C.P. Cross
Wessex Regional Genetics Laboratory
Salisbury NHS Foundation Trust
Salisbury SP2 8BJ, UK

Tel:  +(44) 1722 429080  
Fax:  +(44) 1722 331531  
email:  ncpc@soton.ac.uk
Abstract

Determining the underlying cause of persistent eosinophilia is important for effective clinical management but remains a diagnostic challenge in many cases. We identified STAT5B N642H, an established oncogenic mutation, in 27/1715 (1.6%) cases referred for investigation of eosinophilia. Of the 27 mutated cases, a working diagnosis of hypereosinophilic syndrome (HES; n=7) or a myeloid neoplasm with eosinophilia (n=20) had been made prior to detection of STAT5B N642H. Myeloid panel analysis identified a median of 2 additional mutated genes (range 0-4) with 4 cases having STAT5B N642H as a sole abnormality. STAT5B N642H was absent in cultured T cells of 4/4 positive cases. Individuals with SF3B1 mutations (9/27; 33%) or STAT5B N642H as a sole abnormality had a markedly better overall survival compared to cases with other additional mutations (median 65 months versus 14 months; hazard ratio = 8.1; P<0.001). The overall survival of STAT5B mutated HES cases was only 30 months, suggesting that these cases should be reclassified as chronic eosinophilic leukemia, not otherwise specified (CEL-NOS). The finding of STAT5B N642H as a recurrent mutation in myeloid neoplasia with eosinophilia provides a new diagnostic and prognostic marker as well as a potential target for therapy.

Running title: STAT5B N642H in eosinophilia

Keywords: STAT5B, eosinophilia, myeloid neoplasms
Introduction

Eosinophilia, defined as an elevation of the peripheral blood (PB) eosinophil count above 0.5×10^9/L, is conventionally divided into three main categories: primary, secondary (reactive) and idiopathic. Primary eosinophilia is a clonal hematologic disorder in which the eosinophils form part of the neoplastic clone. Secondary, non-clonal eosinophilia may be driven by a wide range of underlying conditions including allergic disorders, autoimmunity, infectious diseases, lymphoproliferative disorders, solid tumors, drug reactions and other conditions. Idiopathic eosinophilia is a diagnosis of exclusion when no primary or secondary cause can be identified. Clonal eosinophilia is seen in the context of a myeloid neoplasm and particularly the World Health Organisation (WHO)-defined entities ‘chronic eosinophilic leukemia, not otherwise specified’ (CEL-NOS) and ‘myeloid and lymphoid neoplasms with rearrangement of PDGFRα, PDGFRβ or FGFR1 or with PCM1-JAK2, ETV6-JAK2 or BCR-JAK2’ (MLN-eo). Clonal eosinophilia may also be associated with other WHO subtypes of myeloproliferative neoplasm (MPN-eo) or myelodysplastic/myeloproliferative neoplasm (MDS/MPN-eo). Identifying the underlying cause of eosinophilia is important for patient management but can be challenging in the absence of an overt myeloid neoplasm or discernible secondary cause. Primary eosinophilia is strongly associated with constitutively activated tyrosine kinase (TK) signalling, and to date more than 70 TK fusion genes have been identified in myeloid neoplasms as a consequence of reciprocal translocations or other genomic rearrangements. Identification of these fusions usually confirms a specific diagnosis and is an indication for targeted therapy. Notably, imatinib induces rapid and durable complete clinical, hematologic and molecular remissions in >90% of patients with a PDGFRα or PDGFRβ fusion gene, conferring excellent progression-free and overall survival (OS). Fusions involving JAK2, FGFR1 or FLT3 are associated with a more aggressive clinical course and may be responsive to other small molecule inhibitors. Some TK-fusion negative eosinophilia cases test positive for KIT D816V or JAK2 V617F, whereas others have mutations in a range of genes associated with myeloid neoplasms such as TET2, ASXL1, EZH2 or SETBP1. It has been suggested that a rapid and durable response to corticosteroids...
is uncommon in cases with primary eosinophilia and instead points towards a diagnosis of secondary eosinophilia, if that is not already apparent.\textsuperscript{12}

In this study we have used genomic approaches to focus on the identification of novel somatic abnormalities in patients with suspected primary eosinophilia. We have identified a recurrent somatically acquired point mutation in \textit{STAT5B} leading to an N642H substitution in several cases. This mutation is known to activate \textit{STAT5B} but was previously thought to be restricted to lymphoproliferative disorders.

Methods

\textit{Patients and samples.} Our study included samples from individuals referred for routine diagnostic analysis of persistent unexplained eosinophilia and/or patients diagnosed with MPN or MDS/MPN with eosinophilia according to standard morphologic, hematologic and laboratory criteria\textsuperscript{2}. Cases that tested positive for \textit{FIP1L1-PDGFR\alpha} or other recognised TK fusion genes were excluded. The study design adhered to the tenets of the Declaration of Helsinki and was approved by the National Research Ethics Service (UK) Committee South West and the institutional review board of the Medical Faculty of Mannheim, Heidelberg University as part of the ‘German Registry on Disorders of Eosinophils and Mast Cells’. DNA or RNA was extracted from total PB or bone marrow (BM) leukocytes using standard procedures. T-cells were selected using CD3 microbeads (Miltenyi Biotech, Cologne, Germany) and stimulated to divide by co-cultivation with CD2/CD3/CD28 in the presence of interleukin-2 (T-cell activation/expansion kit; Miltenyi Biotech).

\textit{RNAseq analysis.} RNAseq data was analysed for point mutations using an in house pipeline. Briefly, raw RNAseq data in fastq format were aligned to the reference genome (human\textunderscore g1k\textunderscore v37) using STAR aligner and a two-step process. In the first step, splice junctions were detected by an initial alignment of the fastq files. In the second step, final alignments were determined using the splice junctions as a guide.\textsuperscript{13} After alignment the bam files were sorted, de-duplicated and read groups were added using Picard (http://broadinstitute.github.io/picard). In preparation for variant calling, GATK was used to hard-clip intronic sequences (SplitNCigarReads), reassign mapping qualities
Variants were called using GATK HaplotypeCaller and ignoring soft clipped bases, which minimises false positive and false negative calls. A final set of high quality variants in variant call format (VCF) was produced by selecting variants with a phred-scaled confidence threshold of 20 and excluding variants with low depth corrected quality scores (QD<2), significant strand bias (FS>30) or SNP clusters of 3 or more SNPs within 35 base pairs. ANNOVAR was used to annotate variants with respect to genes, databases of known polymorphisms (Exome Aggregation Consortium, 1000 genomes, exome sequencing project, dbSNP), and pathogenicity estimates (avsift, polyphen etc). A shortlist of potentially relevant variants was created by applying additional filters which excluded: non-coding variants, synonymous SNVs and variants with an alternate allele frequency greater than 1% in public databases of known variation: 1000 genomes (1000G; www.ncbi.nlm.nih.gov/variation/tools/1000genomes/), exome sequencing project (ESP; http://evs.gs.washington.edu/EVS/) and exome aggregation consortium (ExAC; http://exac.broadinstitute.org/). Variants passing these filters were further highlighted if they were: located in a manually curated list of candidate genes (n=581 genes), annotated as pathogenic in Clinvar, overlapped with a somatic mutation in hematopoietic and lymphoid tissue in COSMIC (https://cancer.sanger.ac.uk/cosmic) or were recurrently mutated.

Amplification-refractory mutation system (ARMS) test for STAT5B N642H. A tetra-primer ARMS assay was designed using http://primer1.soton.ac.uk/primer1.html with inner primers to specifically amplify the normal and mutant STAT5B alleles and outer primers to produce a positive control STAT5B band for each reaction. PCR primers were: forward outer (FO), 5’-CGATCAGGAAACACGTAGATAAGGTAATT-3; reverse outer (RO), 5’-AAATGGAGATTTCTATTGGAGCCATTAT-3; forward wild-type specific (Fwt), 5’-TCTCTGGTGGTAAAAGGCATCAGGTT-3; reverse mutant specific (Rmt), 5’-TTATTGATCTAGAGGAAAGAATGTTTTAGC-3. The FO and RO were used at a final concentration of 0.5μM whereas the inner primers Fwt and Rmt were used at 2.5μM. Amplification reactions were performed using AmpliTaq Gold DNA polymerase, 25ng genomic DNA at an annealing temperature of 60°C for 35 cycles. Cases showing mutant
bands after agarose gel electrophoresis were confirmed by independent amplification followed by Sanger sequencing.

*Other mutational analysis.* Sanger sequencing was used to screen for other STAT gene mutations using primers designed to amplify from cDNA: STAT3 exons 20-22 (5' -

GGGCCCATCTTGAGCACTAAG-3' and 5'-CACAGATAAACCTGTCCTCAGG-3'); STAT5A exons 16-17 (5'-GGACCTTCTGTGGCTTT-3' and 5'-GGCGGTCAGGAAACACATAG-3') and STAT5B exons 16-17 (5'-AGTGACTCAGAAATGGCCG-3' and 5'-GGCCTGGTCCATGTACGT-3'). ARMS assays were designed for JAK3 V722I and SOCS1 Q201H. For JAK3 V722I a standard tetra primer assay was designed using PCR primers: forward outer (FO), 5'-

CAATAGACCCACCAATCTCCCCAGAC-3'; reverse outer (RO), 5'-

GCAAGGAAGTGGATCCCTGATCCCCACTT-3'; reverse wild-type-specific (Rwt), 5'-

GCCACGTCTGGGAAATGGTGTGAG-3'; forward mutant specific (Fmt), 5'-

ATCCAGGGCACTGTGGAGG-3'. FO, RO, Rwt and Fmt primers were all used at a final concentration of 0.5μM in PCR reactions using AmpliTaq Gold DNA polymerase, 25ng genomic DNA an annealing temperature of 68°C for 35 cycles. Design of a tetra primer assay for SOCS1 Q201H proved to be difficult and therefore we used a simplified mutation-specific assay with primers forward (F) 5'-CCAGGAGGGGGAGCCCCACCAATCTCCCCAGAC-3' and reverse mutant specific (Rmt), 5'-CGCGACTACCTGAGCTTCTCCCCCTTTGTGAC-3' at an annealing temperature of 70°C for 35 cycles. All ARMS assays were validated with control samples, and positive/negative controls were included on each run. The Illumina Trusight Myeloid Sequencing Panel (Illumina, San Diego, CA) was used to screen STAT5B mutated samples for additional somatic mutations. Samples were processed according to the manufacturer’s protocol, run on an Illumina Miseq and results interpreted using Alissa Interpret (Agilent, Cheadle, UK) using a variant allele frequency (vaf) cut off of ≥5%.

*Colony growth and sequencing.* Mononuclear cells from cryopreserved or fresh primary cells were grown in methylcellulose with cytokines, without erythropoietin (Methocult H4035 Optimum, no EPO, Stemcell Technologies (Vancouver, BC, Canada). DNA was prepared for sequencing using the PicoPlex whole genome amplification kit (Rubicon Genomics Inc., Ann Arbor, MI). Colonies were first plucked into 80 μl phosphate buffered saline, spun down and resuspended in 2.5 μl of PicoPlex cell extraction buffer followed by
DNA extraction and amplification according to the manufacturer’s instructions. The DNA was then cleaned up using the QIAquick PCR purification kit (Qiagen, Hilden Germany) prior to amplification and Sanger sequencing using primers designed to target specific mutations.

Statistical analysis. OS probabilities were estimated using the Kaplan-Meier method and compared by the log-rank (Mantel-Cox) test using SPSS v25 (IBM Corporation, Armonk, NY, USA). OS was defined as the time between diagnosis and the date of death or last contact.

Results

STAT5B N642H mutation identified by RNAseq. We previously described RNA-seq analysis to search for cryptic fusion genes in 14 patients with MPN-eo or idiopathic hypereosinophilia with a normal karyotype. Reanalysis of these data for possible point mutations identified 10 candidate variants in 6 cases (Table 1). No variants were seen in 8 cases, including the 2 previously reported to have DIAPH1-PDGFRB or ZMYM2-FLT3 fusions. Two of the 6 cases had known myeloid driver mutations: JAK2 V617F in case E166 and SF3B1 K666N in case E11076. Of the other variants, STAT5B N642H (NM_012448: c.A1924C) was seen in two cases (E11076 and E11493) and stood out as this is a known driver mutation in lymphoproliferative disorders. Furthermore this mutation was recently reported in the T-cells and other cell lineages of two young children with eosinophilia.

Identification of additional cases with STAT5B mutations. We used an ARMS assay to rapidly screen additional retrospective cases for STAT5B N642H (Figure 1a). In an initial screen we found that 2/30 cases with suspected MPN-eo tested positive compared to 0/74 cases with other chronic myeloid neoplasms. Subsequent screening focused only on cases referred for investigation of eosinophilia: of 1671 cases screened, 23 further cases tested positive. Overall, therefore, we found that 27/1715 (1.6%) eosinophilia patients harbored STAT5B N642H. For 25/26 cases confirmed by Sanger sequencing, the mutation specific peak was at a similar height or somewhat lower than the wild type peak, suggesting that the majority of leukocytes were heterozygous for STAT5B N642H. In one sample (E492) only the STAT5B mutant peak was seen suggesting the dominance of a homozygous or hemizygous mutant clone (Supplementary Figure 1). In 4/4 cases tested the STAT5B mutant clone was not
detected in cultured T-cells confirming that it was acquired somatically (Supplementary Figure 2).

*Other STAT mutations.* Other somatic STAT5B and STAT3 mutations have been identified in lymphoproliferative disorders. To test if these mutations might also be associated with persistent eosinophilia, we sequenced known mutation hotspot regions of STAT3 and STAT5B in 153 cases but no variants were seen apart from STAT5B N642H. No human mutations have been reported in STAT5A but due to the high homology with STAT5B we also sequenced part of this gene, but again no variants were found (149 cases).

Additional mutations in STAT5B-mutated cases. We screened all 27 STAT5B N642H cases for mutations in other genes associated with myeloid neoplasms (Figure 1B). Overall, mutations were seen in a median of two additional genes (range 0-4), with 4 cases showing no additional variants (Group 1). Of note, 9 cases had SF3B1 mutations (Group 2) of which 4 were detected as sole additional abnormalities, 3 were seen in combination with single mutations in TET2 or DNMT3A and 2 had additional mutations in genes encoding signalling or splicing components. Fourteen cases did not have SF3B1 mutations but instead had a diverse range of epigenetic, signalling, transcription factor and other splicing mutations with no clear patterns emerging in terms of co-mutated genes (Group 3). All additional mutations are detailed in Supplementary Table 1.

Clinical features associated with STAT5B N642H and other mutations. The clinical features for all cases are summarised on Table 2. There was a preponderance of males (19 males, 8 females) and the median age was 70 (range 7-89; n=27) with a median eosinophil count at presentation of 6x10⁹/L (range 0.5-27; n=26). In most cases eosinophilia was apparent at diagnosis but in 3 cases this was acquired during the course of myelodysplastic syndrome (MDS; cases E2594 and E11837) or MPN (case E13661). Basophilia was noted in several cases. A working diagnosis of idiopathic hypereosinophilic syndrome (HES) had been made in 7 cases prior to the finding of STAT5B N642H and CEL-NOS in 2 cases, supported by the finding of an additional chromosome 8 by cytogenetic analysis in both cases. The Kaplan-Meier estimate for OS of STAT5B-mutated HES cases was only 30 months, which is very short compared to published series¹¹, and suggests that these cases should be reclassified as
CEL-NOS\textsuperscript{10}. The remaining 18 cases had been diagnosed with another myeloid neoplasm, most commonly a subtype of MDS/MPN (n=11). The PB and BM from a representative case with CEL-NOS and \textit{STAT5B} N642H as a sole abnormality is shown on Figure 2.

There was no obvious correlation between the clinical diagnosis and the molecular classification described above. For example the 7 cases initially diagnosed with HES were split between molecular group 1 (n=1), group 2 (n=2) and group 3 (n=4) and the 9 cases with \textit{SF3B1} mutations had been diagnosed with 7 different entities, including HES (n=2) and 6 WHO-defined subtypes of myeloid neoplasms. There was, however, a clear correlation between molecular features and outcome. Focusing on patients with eosinophilia at diagnosis, the OS for cases in group 3 was markedly inferior to cases in groups 1 and 2 [median 14 months versus 65 months; hazard ratio (HR) = 8.1 (95% CI: 1.9 - 23); P<0.0004, Figure 3A]. By contrast age, gender and white cell count were not significantly associated with OS in this relatively small group, whereas eosinophil count approached significance (≤ 6x10\textsuperscript{9}/L, 49 months versus > 6x10\textsuperscript{9}/L, 17 months; P=0.06). No significant effect of mutation number was seen when all mutations were considered but, interestingly, when genes strongly associated with age-related clonal hematopoiesis (i.e. \textit{DNMT3A} and \textit{TET2}) were excluded from the analysis, the OS for cases with mutations in 2 or more additional genes was significantly worse than that of cases with 0 or 1 additional mutations [median 18 months versus 50 months; HR = 6.5 (95% CI: 2.1 – 30) P=0.001, Figure 3B].

As for therapy, 5 patients were treated with imatinib for at least 2 months but none responded. One patient with \textit{KIT} D816V-positive systemic mastocytosis with associated hematologic neoplasm (SM-AHN) responded to midostaurin but the response was lost after 9 months. Some patients showed clinical improvement and reduction in eosinophil counts with corticosteroids but in most cases this was partial and/or transient.

\textbf{Clonal hierarchy.} To understand if acquisition of \textit{STAT5B} N642H is an early or late event in the development of myeloid neoplasia, we genotyped myeloid colonies grown from 4 cases with multiple mutations. For case E10926, 6 mutations in 5 genes had been identified by bulk analysis. By colony analysis we found a major clone with 5 mutations that included \textit{STAT5B} N642H and \textit{KRAS} V14I but not \textit{KIT} D816V, and a minor clone that included \textit{KIT}
D816V but not STAT5B N642H or KRAS V14I. All colonies tested positive for SF3B1 and two independent TET2 mutations and thus branching evolution can be inferred with STAT5B N642H acquired in a later subclone. For 2 of the other cases linear evolution was apparent with STAT5B and ASXL1 mutations acquired as late events. The fourth case was uninformative with all colonies positive for the 3 mutations detected on bulk analysis (Figure 4A). For two cases with increasing eosinophil counts, the STAT5B N642H vaf, as estimated by Sanger sequencing, increased over time (Figure 4B).

Assessment of other variants identified by RNAseq. Other recurrent mutations identified by RNAseq were also considered. Both PRF1 N252S and SOCS1 Q210H were identified in two individuals (Table 1). PRF1 encodes perforin 1, a gene associated with familial hemophagocytic lymphohistiocytosis. However the N252S variant has been shown to be a non-functional rare single nucleotide polymorphism (SNP) and was therefore not considered further. SOCS1 encodes suppressor of cytokine signalling 1, which takes part in a negative feedback loop to attenuate cytokine signalling and therefore has obvious potential relevance to myeloid neoplasia. The status of the Q210H variant is unclear: it is seen at a frequency of 0.2-0.6% in the 1000G, ESP and ExAC control datasets, as are some known somatic driver mutations such as JAK2 V617F. For one case we tested buccal cells and found the variant was constitutional (data not shown). To test the possibility that Q210H might be more widespread in myeloid neoplasia we tested cases referred for investigation of eosinophilia (n=151), mastocytosis (n=75), MDS/MPN (n=54) and healthy controls (n=88). Just 3 (1%) cases tested positive (one with mastocytosis, 2 with MDS/MPN) and none of the controls (P=0.99; Fisher’s exact test) suggesting that this variant is an irrelevant, rare polymorphism that happened to be present in 2 of the cases analysed by RNAseq. JAK3 V722I was seen in a single case and has been reported previously as an activating mutation but is also present at 0.4-1% in control datasets and its pathogenicity has been disputed. We found that 4/307 cases with suspected myeloid neoplasia tested positive for JAK3 V722I (one of which also had the mutation in cultured T-cells) as well as 3/88 normal controls (P=0.19), again suggesting it is an irrelevant polymorphism. Similarly, RPS19 T55M is also believed to be a polymorphism.

11
Discussion

STAT5 is a key component of cytokine-induced signal transduction cascades, and a critical downstream mediator of transformation by oncogenic tyrosine kinases such as BCR-ABL1, JAK2 V617F, FLT3-ITD and ZNF198-FGFR1. STAT5 is encoded by two different genes, STAT5A and STAT5B, located closely together at chromosome 17q11.2 that encode proteins with >90% amino acid identity and largely redundant functions. Nevertheless, targeted disruption of STAT5A and STAT5B gives rise to distinct phenotypes in mice and only STAT5B has been reported to be a target of mutations in cancer.

STAT5A N642H was initially identified as a constitutively activating mutation in a random ex vivo mutagenesis screen. The corresponding mutation in STAT5B was subsequently identified as a somatically acquired driver mutation in 1-37% of patients with various lymphoid neoplasms including large granular lymphocytic leukemia, paediatric T-cell acute lymphoblastic leukemia, T-cell prolymphocytic leukemia, γδ-T-cell lymphoma and two cases of lymphocyte-driven early onset nonclonal eosinophilia with urticaria, dermatitis and other features. The most frequent mutation seen in these T-cell disorders is STAT5B N642H, but other mutations are also seen, specifically STAT5B Y665F and STAT3 Y640F, N647I and D661V/Y. Isolated reports have identified single cases of myeloid neoplasms, specifically chronic neutrophilic leukemia and MLN-eo, that tested positive for STAT5B N642H as well as 2 cases that developed clonal hematopoiesis following aplastic anemia. Our findings are the first to identify STAT5B N642H as a recurrent mutation in myeloid neoplasms with eosinophilia.

We demonstrated the absence of STAT5B N642H in cultured T-cells from 4 cases, one of whom (case E12614) was in molecular group 1, i.e. had no additional mutations. All 4 group 1 cases, and indeed most cases in this study, had STAT5B vafs which suggested the great majority of leukocytes were heterozygously mutated. Although we cannot exclude the possibility that some T-cell subsets might be part of the mutant clone, our data strongly suggest that STAT5B N642H drives primary eosinophilia irrespective of the presence or absence of additional mutations.
Cytokine stimulation results in phosphorylation of STAT5B by receptor or non-receptor tyrosine kinases. Dimerization of tyrosine phosphorylated (pY) STAT5B is mediated by trans-SH2 domain/phosphotyrosine binding, and the dimers then translocate to the nucleus and activate transcription of target genes. N642 is located within the SH2 domain of STAT5B, close to the phosphotyrosine-binding loop. Rather than being constitutively active, the N642H mutant shows prolonged pY-STAT5 levels upon cytokine stimulation due to greatly enhanced stability of N642H homodimers. Transgenic expression of STAT5B N642H under the control of the Vav1 promoter (which is believed to be active in all hematopoietic cell types, including stem cells) resulted in transplantable CD8+ T-cell neoplasia. Although eosinophilia was not noted, there was a mild expansion of mature and immature myeloid lineage cells. The apparent absence of myeloid neoplasia with eosinophilia in this mouse model may be the result of a number of factors, including the precise expression of STAT5B, the effect of other somatic mutations, the cell type in which STAT5B N642H arises and the rapid onset of T-cell neoplasia.

We observed a striking influence of additional mutations on patients with STAT5B N642H that mirrors established associations in related myeloid disorders, particularly the relatively good prognostic influence of SF3B1 mutations and the adverse influence of multiple mutations, both of which have been described in MDS. In addition, only one case with an SF3B1 mutation tested positive for one or more mutations in SRSF2, ASXL1 and RUNX1 (S/A/R), known to be an adverse prognostic factor in advanced SM, whereas 10/14 cases in group 3 were S/A/R positive (P=0.009; Fisher’s exact test).

The combination of JAK2 V617F and SF3B1 mutations is associated with MDS/MPN with ring sideroblasts and thrombocytosis. In our series, the working diagnosis of patients with STAT5B N642H and SF3B1 mutations ranged from HES to MDS/MPN (Table 2). Stored material from most cases was not available for central morphological review but we envisage that prospective analysis of new cases will help to define more accurately the features associated with STAT5B N642H in the presence or absence of SF3B1 mutations.

Previous studies have shown that STAT5A is required for eosinophil differentiation of cord blood-derived CD34+ cells. Our findings, however, suggest that STAT5B N642H may be a
driver of eosinophilia. First, in most of our cases eosinophilia was apparent at diagnosis and the level of STAT5B N642H as assessed by Sanger sequencing indicated that the mutation was present in the majority of cells. In 4 of these cases STAT5B N642H was detected as a sole abnormality, although we cannot exclude the possibility of mutations in genes not covered by the myeloid panel. Second, analysis of 2 cases with increasing eosinophil counts showed an increase in the STAT5B N642H vaf. Three additional cases acquired eosinophilia during the course of their disease but unfortunately samples were not available for analysis from the pre-eosinophilia phase.

The finding that STAT5B N642H shows prolonged activation following cytokine signalling suggests that targeting upstream tyrosine kinases may ameliorate the activity of mutant STAT5B. Indeed, the T-cell neoplasms induced by transgenic STAT5B N642H was markedly supressed by JAK1/2 inhibition\textsuperscript{35}. Furthermore, a small molecule inhibitor of STAT5B dimerization has been shown to inhibit the growth of FLT3-ITD positive AML cells\textsuperscript{41}. Further studies will be required to determine if myeloid disorders associated with STAT5B N642H are targetable with small molecule inhibitors.

Acknowledgements

This study was funded by Bloodwise Specialist Programme Grant no. 13002 to NCPC, WJT and AC. We thank Professor Satu Mustjoki, Helsinki for providing control samples with STAT5B mutations.

Conflicts of interest

None of the authors have any relevant conflicts of interest or disclosures


Figure legends

Figure 1. (A) ARMS assay for STAT5B N642H. Heterozygotes have 3 bands - control band (243bp), wild type (wt) specific band (166bp) and mutant allele specific band (124bp). Using serial dilutions of DNA from a mutated case we estimated that the assay can detect STAT5B N642H at a variant allele frequency of 10%; (B) summary of mutations for all 27 STAT5B N642H mutated cases as determined by ARMS PCR, Sanger sequencing and myeloid panel analysis. Group 1 are cases with no additional mutated genes, group 2 have additional mutated genes that include SF3B1 and group 3 have additional mutated genes that do not include SF3B1. ARCH = genes strongly associated with age-related clonal hematopoiesis. Full details of additional mutations are given in Supplementary Table 1.

Figure 2. Peripheral blood film (A), bone marrow smear (B) and bone marrow trephine at x20 (C) and x100 (D). The blood film showed 40% neutrophils, 28% eosinophils, 4% basophils, 12% lymphocytes and 7% monocytes. The eosinophils were morphologically close to normal with only a very minor degree of vacuolation and hypogranularity but there was an increase in non-lobated forms. Bone marrow cellularity and megakaryocyte numbers are increased. Increased cellularity is due to an increase in all three granulocyte lineages (neutrophils, basophils and eosinophils). The trephine biopsy sections show hypercellularity, disorganisation and an increase in cells of neutrophil and eosinophil lineages. Reticulin was not increased.

Figure 3. Kaplan-Meier plots showing overall survival (OS) estimates for the STAT5B N642H mutated cases who had eosinophilia at diagnosis and had follow up data available (n=23). (A) Comparison of the 3 molecular groups shows that cases in group 3 (those with additional mutated genes that do not include SF3B1) have an inferior OS compared to all other cases (median 14 months versus 65 months; P<0.0004). (B) OS for cases with mutations in 2 or more additional genes (excluding DNMT3A and TET2) was significantly worse than that for cases with 0 or 1 additional mutations (median 18 months versus 50 months; P=0.001). Of the 9 cases with mutations in ≥2 additional genes, 8 were in group 3.
Figure 4. (A) Clonal hierarchy for 4 patients. The numbers indicate the number of colonies with that genotype, e.g. for case E11825, 5 colonies were mutant for TET2 and SF3B1 but not STAT5B and 9 colonies were mutant for all 3 genes. Light grey indicates an inferred ancestral clone. (B) Sequential analysis of two cases with increasing eosinophil counts (values indicated are ×10⁹/L). Case E11825 had an elevated eosinophil count at diagnosis, with eosinophilia increasing over time co-incident with an increase in STAT5B N642H vaf. Case E13661 was diagnosed initially with PV in 1988 and STAT5B N642H was not detected in a sample taken 6 years later. Slightly elevated eosinophil counts were first noted in 2014, and a sample from this time was positive for STAT5B N642H. The mutant vaf and eosinophil counts increased over the next 4 years.
<table>
<thead>
<tr>
<th>Case</th>
<th>Gene</th>
<th>Refseq</th>
<th>cDNA</th>
<th>protein</th>
<th>vaf</th>
</tr>
</thead>
<tbody>
<tr>
<td>E166</td>
<td>JAK2</td>
<td>NM_004972</td>
<td>c.G1849T</td>
<td>p.V617F</td>
<td>1.00</td>
</tr>
<tr>
<td>E166</td>
<td>SOCS1</td>
<td>NM_003745</td>
<td>c.G630C</td>
<td>p.Q210H</td>
<td>0.48</td>
</tr>
<tr>
<td>E11072</td>
<td>RPS19</td>
<td>NM_001022</td>
<td>c.C164T</td>
<td>p.T55M</td>
<td>0.48</td>
</tr>
<tr>
<td>E11072</td>
<td>SOCS1</td>
<td>NM_003745</td>
<td>c.G630C</td>
<td>p.Q210H</td>
<td>0.86</td>
</tr>
<tr>
<td>E11075</td>
<td>PRF1</td>
<td>NM_001083116</td>
<td>c.A755G</td>
<td>p.N252S</td>
<td>0.39</td>
</tr>
<tr>
<td>E11076</td>
<td>SF3B1</td>
<td>NM_012433</td>
<td>c.G1998T</td>
<td>p.K666N</td>
<td>0.32</td>
</tr>
<tr>
<td>E11076</td>
<td>PRF1</td>
<td>NM_001083116</td>
<td>c.A755G</td>
<td>p.N252S</td>
<td>0.67</td>
</tr>
<tr>
<td>E11076</td>
<td>STAT5B</td>
<td>NM_012448</td>
<td>c.A1924C</td>
<td>p.N642H</td>
<td>0.43</td>
</tr>
<tr>
<td>E11493</td>
<td>STAT5B</td>
<td>NM_012448</td>
<td>c.A1924C</td>
<td>p.N642H</td>
<td>0.47</td>
</tr>
<tr>
<td>E11500</td>
<td>JAK3</td>
<td>NM_000215</td>
<td>c.G2164A</td>
<td>p.V722I</td>
<td>0.56</td>
</tr>
</tbody>
</table>
Table 2. Clinical summary of \textit{STAT5B} N642H mutant cases

<table>
<thead>
<tr>
<th>Case</th>
<th>Sex</th>
<th>Age (y)</th>
<th>Status (m)</th>
<th>wcc</th>
<th>plt</th>
<th>Hb</th>
<th>eos</th>
<th>Initial diagnosis</th>
<th>MG</th>
</tr>
</thead>
<tbody>
<tr>
<td>E143</td>
<td>M</td>
<td>69</td>
<td>D (107)</td>
<td>16</td>
<td>280</td>
<td>100</td>
<td>1.5</td>
<td>aCML</td>
<td>2</td>
</tr>
<tr>
<td>E228</td>
<td>M</td>
<td>65</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>HES</td>
<td>2</td>
</tr>
<tr>
<td>E492</td>
<td>F</td>
<td>7</td>
<td>D (13)</td>
<td>30</td>
<td>343</td>
<td>147</td>
<td>11</td>
<td>Myeloid sarcoma</td>
<td>3</td>
</tr>
<tr>
<td>E1237</td>
<td>M</td>
<td>89</td>
<td>D (12)</td>
<td>15</td>
<td>488</td>
<td>148</td>
<td>6</td>
<td>MPN-U</td>
<td>3</td>
</tr>
<tr>
<td>E2323</td>
<td>M</td>
<td>70</td>
<td>D (12)</td>
<td>61</td>
<td>107</td>
<td>101</td>
<td>10</td>
<td>CMLM</td>
<td>3</td>
</tr>
<tr>
<td>E2335</td>
<td>M</td>
<td>63</td>
<td>D (9)</td>
<td>28</td>
<td>113</td>
<td>158</td>
<td>0.7</td>
<td>MDS/MPN</td>
<td>3</td>
</tr>
<tr>
<td>E2594</td>
<td>M</td>
<td>58</td>
<td>D (14)</td>
<td>14</td>
<td>128</td>
<td>81</td>
<td>4</td>
<td>MDS-SLD; eosinophilia at 3y</td>
<td>2</td>
</tr>
<tr>
<td>E2810</td>
<td>M</td>
<td>81</td>
<td>D (7)</td>
<td>18</td>
<td>409</td>
<td>94</td>
<td>10</td>
<td>MPN-U</td>
<td>2</td>
</tr>
<tr>
<td>E2935</td>
<td>F</td>
<td>10</td>
<td>D (48)</td>
<td>47</td>
<td>157</td>
<td>110</td>
<td>27</td>
<td>HES</td>
<td>1</td>
</tr>
<tr>
<td>E3065</td>
<td>F</td>
<td>82</td>
<td>D (28)</td>
<td>25</td>
<td>150</td>
<td>100</td>
<td>2</td>
<td>CMML + MM</td>
<td>1</td>
</tr>
<tr>
<td>E3321</td>
<td>M</td>
<td>86</td>
<td>D (45)</td>
<td>21</td>
<td>n/a</td>
<td>152</td>
<td>6</td>
<td>HES</td>
<td>2</td>
</tr>
<tr>
<td>E4349</td>
<td>M</td>
<td>61</td>
<td>D (12)</td>
<td>15</td>
<td>16</td>
<td>132</td>
<td>5</td>
<td>CMML</td>
<td>3</td>
</tr>
<tr>
<td>E4735</td>
<td>M</td>
<td>68</td>
<td>D (31)</td>
<td>12</td>
<td>106</td>
<td>111</td>
<td>5</td>
<td>HES</td>
<td>3</td>
</tr>
<tr>
<td>E10926</td>
<td>M</td>
<td>70</td>
<td>D (16)</td>
<td>142</td>
<td>111</td>
<td>73</td>
<td>1.4</td>
<td>SM-AHN</td>
<td>2</td>
</tr>
<tr>
<td>E11076</td>
<td>M</td>
<td>72</td>
<td>A (68)</td>
<td>21</td>
<td>556</td>
<td>149</td>
<td>6</td>
<td>MDS/MPN-U</td>
<td>2</td>
</tr>
<tr>
<td>E11493</td>
<td>M</td>
<td>76</td>
<td>D (23)</td>
<td>42</td>
<td>275</td>
<td>92</td>
<td>19</td>
<td>MDS/MPN-U</td>
<td>3</td>
</tr>
<tr>
<td>E11825</td>
<td>F</td>
<td>61</td>
<td>A (61)</td>
<td>9</td>
<td>438</td>
<td>108</td>
<td>1.2</td>
<td>MDS/MPN-RS-T</td>
<td>3*</td>
</tr>
<tr>
<td>E11837</td>
<td>M</td>
<td>72</td>
<td>D (14)</td>
<td>11</td>
<td>120</td>
<td>156</td>
<td>7</td>
<td>MDS-MLD; eosinophilia at 3y</td>
<td>1</td>
</tr>
<tr>
<td>E12398</td>
<td>M</td>
<td>77</td>
<td>D (15)</td>
<td>24</td>
<td>34</td>
<td>140</td>
<td>7</td>
<td>HES</td>
<td>3*</td>
</tr>
<tr>
<td>E12433</td>
<td>F</td>
<td>51</td>
<td>A (35)</td>
<td>19</td>
<td>151</td>
<td>90</td>
<td>3</td>
<td>MDS/MPN-U</td>
<td>2</td>
</tr>
<tr>
<td>E12601</td>
<td>F</td>
<td>64</td>
<td>D (14)</td>
<td>28</td>
<td>267</td>
<td>116</td>
<td>10</td>
<td>HES</td>
<td>3</td>
</tr>
<tr>
<td>E12614</td>
<td>F</td>
<td>15</td>
<td>A (48)</td>
<td>28</td>
<td>655</td>
<td>133</td>
<td>4</td>
<td>CEL-NOS</td>
<td>1*</td>
</tr>
<tr>
<td>E12885</td>
<td>F</td>
<td>75</td>
<td>D (23)</td>
<td>24</td>
<td>238</td>
<td>110</td>
<td>3</td>
<td>MDS/MPN</td>
<td>3*</td>
</tr>
<tr>
<td>E13191</td>
<td>M</td>
<td>83</td>
<td>A (9)</td>
<td>29</td>
<td>401</td>
<td>148</td>
<td>13</td>
<td>HES</td>
<td>3</td>
</tr>
<tr>
<td>E13206</td>
<td>M</td>
<td>85</td>
<td>D (4)</td>
<td>39</td>
<td>229</td>
<td>142</td>
<td>20</td>
<td>CEL-NOS</td>
<td>3</td>
</tr>
<tr>
<td>E13493</td>
<td>M</td>
<td>70</td>
<td>D (9)</td>
<td>17</td>
<td>195</td>
<td>65</td>
<td>6</td>
<td>MDS/MPN-U</td>
<td>3</td>
</tr>
<tr>
<td>E13661</td>
<td>M</td>
<td>42</td>
<td>A (60)</td>
<td>12</td>
<td>267</td>
<td>125</td>
<td>0.5</td>
<td>PV; eosinophilia at 25y</td>
<td>3</td>
</tr>
</tbody>
</table>

Sex: male (M) or female (F); Age in years (y) at presentation; Status: dead (D) or alive (A) at specified number of months (m) after presentation or first detection of eosinophilia (E2594, E11837, E13661); white cell count (wcc), platelets (plt) and eosinophils (eos) ×10^9/L at diagnosis or first detection of eosinophilia; hemoglobin (Hb) in g/L; diagnosis: atypical chronic myeloid leukemia (aCML), hypereosinophilic syndrome (HES), myeloproliferative neoplasm, unclassifiable (MPN-U); myelodysplastic/myeloproliferative neoplasm (MDS/MPN), myelodysplastic syndrome with single lineage dysplasia (MDS-SLD), chronic myelomonocytic leukemia (CMML), multiple myeloma (MM), systemic mastocytosis with associated hematological neoplasm (SM-AHN), MDS/MPN, unclassifiable (MDS/MPN-U), MDS/MPN with ring sideroblasts and thrombocytosis (MDS/MPN-RS-T), MDS with multilineage dysplasia (MDS-MLD), chronic eosinophilic leukemia, not otherwise specified (CEL-NOS), polycythemia vera (PV); molecular group (MG): 1 = no additional mutations; 2 = \textit{SF3B1} mutated; 3 = additional mutations but \textit{SF3B1} unmutated; not available (n/a). * indicates the 4 cases for whom \textit{STAT5B} N642H was absent in cultured T-cells.
Figure 1

A

B

<table>
<thead>
<tr>
<th></th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1935</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E13065</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E11837</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E12614</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E143</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E228</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E15094</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E2810</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E3211</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E10926</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E11076</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E11825</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E1492</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E1237</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E1233</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E3476</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E1238</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E3489</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E1493</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E12398</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E12601</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E13191</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E12393</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E13661</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| STAT5B        |         |         |         |
| SF3B1         |         |         |         |
| SRSF2         |         |         |         |
| U2AF1         |         |         |         |
| ZRSR2         |         |         |         |
| TET2          |         |         |         |
| DNMT3A        |         |         |         |
| KIT           |         |         |         |
| NRAS          |         |         |         |
| CBL           |         |         |         |
| KRAS          |         |         |         |
| JAK2          |         |         |         |
| IDH2          |         |         |         |
| ASXL1         |         |         |         |
| EZH2          |         |         |         |
| CUX1          |         |         |         |
| RUNX1         |         |         |         |
Figure 3

A

Overall survival

- Group 1 (n=3)
- Group 2 (n=7)
- Group 3 (n=13)

Time (months)

P<0.001

B

Overall survival

- 0-1 additional genes (n=14)
- ≥2 additional genes (n=9)

Time (months)

P=0.001
Figure 4

A

E10926

\[ \text{SF3B1} \]
\[ \text{TET2}* \]
\[ \text{TET2del} \]

\[ \text{SF3B1} \]
\[ \text{TET2}* \]
\[ \text{TET2del} \]
\[ \text{STAT5B} \]
\[ n=15 \]

\[ \text{SF3B1} \]
\[ \text{TET2} \]
\[ \text{del} \]
\[ \text{KIT} \]
\[ n=2 \]

E11825

\[ \text{TET2} \]
\[ \text{SF3B1} \]
\[ n=5 \]

\[ \text{TET2} \]
\[ \text{SF3B1} \]
\[ \text{STAT5B} \]
\[ n=9 \]

E12398

\[ \text{STAT5B} \]
\[ \text{CUX1} \]
\[ n=14 \]

\[ \text{STAT5B} \]
\[ \text{CUX1} \]
\[ \text{ASXL1} \]
\[ n=10 \]

E12885

\[ \text{STAT5B} \]
\[ \text{ASXL1} \]
\[ \text{U2AF1} \]
\[ n=14 \]

B

E11825

2014; eos = 1.2

2015; eos = 5.8

E13661

1994; PV

2014; eos = 0.4

2018; eos = 3