Capture-based targeted sequencing using a T-cell control in myeloid malignancies and idiopathic cytopenias

Pietka, G., De Lord, C., Matthias, G., Cheung, B., Atwal, S., Furtado, M., Cullis, J., Grey-Davies, L., Narayanan, S., McGregor, A., Kilner, M., Bosworth, J., McMullin, M. F., Coats, T., Parcharidou, A., Cavenagh, J., Byrne, J., Iyengar, S., Mohammed, K., ... Taussig, D. (2024). Capture-based targeted sequencing using a T-cell control in myeloid malignancies and idiopathic cytopenias. *British Journal of Haematology*. Advance online publication. https://doi.org/10.1111/bjh.19377

Published in:
British Journal of Haematology

Document Version:
Publisher's PDF, also known as Version of record

Received: 21 November 2023 | Accepted: 20 February 2024

DOI: 10.1111/bjh.19377

BJHaem
BRITISH JOURNAL OF HAEMATOLOGY

ORIGINAL PAPER

Haematological Malignancy – Biology

Capture-based targeted sequencing using a T-cell control in myeloid malignancies and idiopathic cytopenias

Grzegorz Pietka[1,2] | Corinne De Lord[3,4] | Gwynn Matthias[5] | Betty Cheung[6] | Sangeeta Atwal[7] | Michelle Furtado[8] | Jonathan Cullis[9] | Liz Grey-Davies[10] | Srinivasan Narayanan[11] | Andrew McGregor[12] | Mari Kilner[13] | Jenny Bosworth[3,4] | Mary Frances McMullin[14] | Thomas Coats[15] | Agapi Parcharidou[16] | Jamie Cavenagh[17] | Jenny Byrne[18] | Sunil Iyengar[4] | Kabir Mohammed[19] | Nicholas Cross[20,21] | Mike Hubank[22,23] | Sara Ribeiro[24] | Jamshid Khorashad[24] | Dorte Wren[25] | Simon O'Connor[26] | David Taussig[1,4]

Correspondence
David Taussig, Department of Haematology, Royal Marsden NHS Foundation Trust, London, UK.
Email: david.taussig@icr.ac.uk

Funding information
NIHR Biomedical Research Centre, Royal Marsden NHS Foundation Trust/Institute of Cancer Research, Grant/Award Number: A127

Summary
We report on a study of next-generation sequencing in 257 patients undergoing investigations for cytopenias. We sequenced bone marrow aspirates using a target enrichment panel comprising 82 genes and used T cells from paired blood as a control. One hundred and sixty patients had idiopathic cytopenias, 81 had myeloid malignancies and 16 had lymphoid malignancies or other diagnoses. Forty-seven of the 160 patients with idiopathic cytopenias had evidence of somatic pathogenic variants consistent with clonal cytopenias. Only 39 genes of the 82 tested were mutated in the 241 patients with either idiopathic cytopenias or myeloid neoplasms. We confirm that T cells can be used as a control to distinguish between germline and somatic variants. The use of paired analysis with a T-cell control significantly reduced the time molecular scientists spent reporting compared to unpaired analysis. We identified somatic variants of uncertain significance (VUS) in a higher proportion (24%) of patients with myeloid malignancies or clonal cytopenias compared to less than 2% of patients with non-clonal cytopenias. This suggests that somatic VUS are indicators of a clonal process. Lastly, we show that blood depleted of lymphocytes can be used in place of bone marrow as a source of material for sequencing.

KEYWORDS
germline control, idiopathic cytopenia, myelodysplastic syndrome, next-generation sequencing, T-cell control, variant of unknown significance

INTRODUCTION

Next-generation targeted sequencing (NGS) is an established diagnostic tool for haematological malignancies. NGS is performed by aligning the sequenced reads to the genome reference sequence and reporting discovered differences. One limitation of this approach is the inherent heterogeneity of germline genetics, which results in the inability to unambiguously determine whether any differences detected between tumour and reference are

For Affiliation refer page on 8

germline or somatic in origin. Single nucleotide polymorphisms (SNP) can be filtered out based on their population frequency using resources such as gnomAD or dbSNP, or based on their clinical significance using databases such as ClinVar or COSMIC.[1–5] Once SNPs have been excluded, variants of unknown significance (VUS) remain, defined here as a variant less common than 1 in 10 000 of the population that alters the protein sequence but is not known to be pathogenic or likely pathogenic (P/LP) and is not reported in homozygous carriers. However, tumour-only sequencing can lack precision, with a false-positive rate of 31% reported for somatic variants in solid malignancies in one study and a false-negative rate of 44% for true somatic variants in another.[6,7] One method to improve the distinction between germline and somatic variants is to use a paired control germline sample from the patient (often referred to as 'normal'), which allows identification of tumour-specific acquired variants.

In this paper, we report findings from a paired tumour-normal sequencing study using an in-house-designed targeted capture panel, the primary aim of which was to investigate patients with idiopathic cytopenias (IC). We chose 82 genes for our panel based on genes most commonly mutated in MDS, focusing on those with the most significant prognostic and diagnostic relevance.[8–15] Capture targets for 75 genes were designed to investigate the full coding sequence; five genes were focused on hotspot exons, and two genes were assessed for copy number only (Table S1). All 82 genes had additional probes designed to determine copy number changes.

Bone marrow was used as the primary diagnostic sample. Previous work suggests that T cells do not normally carry the MDS-associated variants present in myeloid lineage cells.[16] Therefore, we opted to use T cells as a control.

Most older adults have leukaemia-associated variants detectable in the blood when high-sensitivity sequencing is employed.[17] However, low variant allele frequency (VAF) clones are not thought to disturb haemopoiesis. Our previous work on the impact of leukaemia on normal haemopoiesis suggests that at least 20% of marrow cells need to be leukaemic to perturb normal haemopoiesis significantly.[18] Consistent with our findings, Buscarlet et al. noted a correlation between variant VAF >10% and cytopenias.[19] Given our focus was on diagnosis in IC, not prognosis in AML/MDS, we selected a VAF threshold of 10% as we aimed for the identification of sufficiently large clones that might cause the cytopenia under investigation.

METHODS

T-cell separation and purity assessment

T-cell separation from peripheral blood was performed using the EasySep™ Human Whole Blood CD3 Positive Selection Kit (StemCell). For additional details, see supplemental methods.

DNA extraction

DNA was extracted using the QIAsymphony DNA Midi Kit (Qiagen) and quantified using Qubit Broad Range and High Sensitivity (Thermofisher) assay kits.

NGS library preparation, capture and sequencing

Next-generation sequencing libraries were prepared using the HyperPlus NGS Library Preparation Kit (Roche) and the KAPA Dual-Indexed Adapter Kit (Roche). Libraries were pooled and captured using the Nimblegen SeqCap Custom Capture Panel. Sequencing was performed using the NextSeq Mid Output 150 cycle kit (Illumina).

NGS sequencing data processing

The bioinformatics pipeline is capable of reporting SNV, CNV and structural variants within the panel scope. At a 5% VAF threshold, the assay is validated with the following specifications: sensitivity of cancer SNV detection >98% [95% CI: 95.94%]; specificity of cancer SNV detection >98% [95% CI: 95.94%]; accuracy of cancer SNV detection >98% [95% CI: 98.21%–100%]; sensitivity of cancer indel detection >91.7% [95% CI: 89.11%]; accuracy of cancer indel detection ≥89% [95% CI: 73.94%–96.89%]. This method was clinically validated in accordance with the ISO15189 standard, and each sequencing run contains a positive and negative control for performance and quality monitoring. BCL files were converted to FASTQ using Illumina bcl2fastq2. The BWA MEM algorithm was used to align sequencing data to the human genome (hg19). Sequencing data QC was performed using Picard. GATK Best Practice pipeline 4.0, designed by Broad Institute, was used for next-generation sequencing variant analysis. Germline variant calling was performed by GATK 4.0, and somatic mutations were detected using Mutect2. Variants were annotated using Oncotator against version-controlled public databases (gnomAD, COSMIC).

Mutation detection

The presence of mutations produced by the analysis pipeline was investigated by manual checking of the BAM files in Integrative Genomics Viewer (Broad Institute). Variants reported by the pipeline underwent a double-blind manual analysis in which a genetic technologist and a clinical scientist performed an assessment according to guidelines.[20–22] The final result was compared by a clinical scientist, who assigned pathogenicity to the identified variants in accordance with the AMP/ASCO/AGCS classification guidelines based on their frequency in the population and annotation in COSMIC, dbSNP and ClinVar.[20–23]

When required, an additional review of published material was performed.

Statistical analysis

Descriptive statistical analysis methods were used to summarise the data using frequency and percentages for categorical data and mean/median and standard deviation or interquartile range (IQR) for continuous data.

Categorical variables were compared using Fisher's exact or Chi square tests, and continuous variables using the Mann–Whitney *U*-test, Student's *t*-test or Wilcoxon test.

All *p*-values were two-tailed; *p*-values <0.05 were considered as statistically significant. For comparison of gene frequencies and VAF between groups, we only selected genes with a variant allele frequency of 10% or higher. ± indicates the standard deviation.

RESULTS

Diagnosis distribution of patients

We prospectively recruited 303 adult patients with cytopenias to the Improving Diagnosis in Idiopathic Cytopenia Using Gene Sequencing trial (NCT03026751) from 2017 to 2021.

Out of 303 consecutively recruited patients, 160 were confirmed to have IC based on standard pathology workup of the bone marrow (aspirate/trephine morphology, cytogenetics, immunophenotyping) and had their bone marrow and T cells sequenced. An additional 97 patients with successful sequencing of T cells and marrow had a diagnosis of a haematological disorder and were divided into two groups; those with myeloid neoplasms (MN), totalling 81 patients (acute myeloid leukaemia (AML) $n = 18$, myelodysplastic syndrome (MDS) $n = 49$, chronic myelomonocytic leukaemia (CMML) $n = 8$ and myelofibrosis $n = 3$, aplastic anaemia $n = 1$, cytogenetically defined clonal cytopenia $n = 1$, systemic mastocytosis $n = 1$), and those with lymphoid/other diagnoses ($n = 16$). The remaining 46 patients failed to meet inclusion/exclusion criteria or had inadequate material for sequencing (consent withdrawal $n = 3$, normal blood count at screening $n = 3$, inadequate standard diagnostic workup $n = 17$ or inadequate material sent for sequencing $n = 23$).

Molecular profile of patients

We generated NGS data from 257 patients (160 with IC and the rest with MN/lymphoid/other diagnoses) using a custom-tailored capture panel and an in-house bioinformatical pipeline validated according to Medical Laboratory Accreditation standard ISO15189. We derived our control DNA from immunomagnetically selected T cells with a median purity of 97.3% (range: 50.7%–99.8%, Figure S1). Tumour diagnostic and control T-cell samples were sequenced to an average depth of 455 reads and 177 reads respectively. Both tumour and germline sequencing results were analysed using the pipeline; however, variant calling was performed on the tumour sample only, and the germline sample was used as a reference during manual variant review using the Integrative Genomics Viewer (IGV) (Figures S2 and S3). Germline variants are expected to occur with a VAF of around 40%–60%; however, they may fall outside this range due to sample quality, target gene or sequencing depth.[24] Any variant present at 36%–60% VAF in both tumour and control samples was classified as germline.

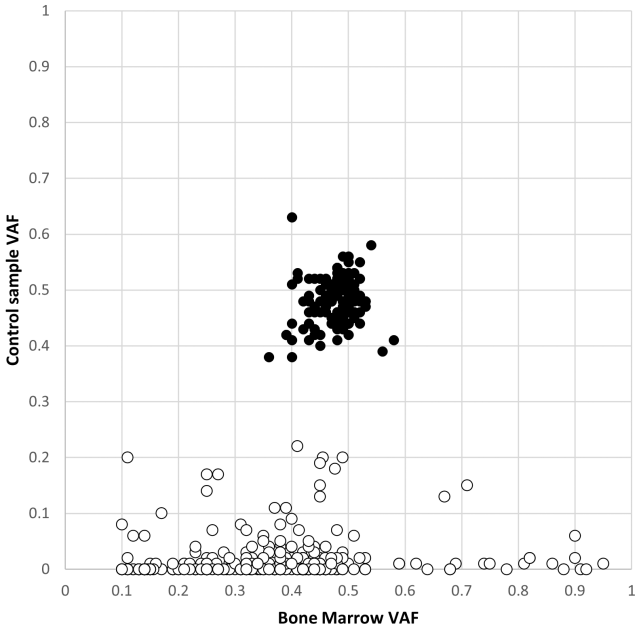

FIGURE 1 VAFs of somatic and germline variants in paired bone marrow and T-cell control samples. There was no significant difference in VAF between the T-cell control and bone marrow samples for germline variants (black), including P/LP variants and VUS ($n = 142$ variants, Wilcoxon test $p = 0.97$). The VAF of somatic variants (white), including P/LP variants and VUS, was significantly higher in BM than in the paired control T-cell samples. ($n = 298$ variants, Wilcoxon test $p < 0.0001$).

We compared the VAF of variants (both P/LP and VUS) in the bone marrow (BM) and paired T-cell control sample. The median VAF of all germline variants (both germline P/LP variants and germline VUS) ($n = 144$ variants in 105 patients) was 48% (range 36%–58%) and 48% (range 38%–63%) in bone marrow and T-cell control samples respectively (Figure 1; Wilcoxon test $p = 0.65$). The median VAF of all somatic variants ($n = 305$ variants in 119 patients) (both somatic pathogenic P/LP variants and somatic VUS) was 38% (range 10%–95%) and 6% (range 1%–22%) in bone marrow and T-cell control samples respectively (Figure 1; Wilcoxon test $p < 0.0001$). Three germline VUS had a VAF of less than 40% (36%–39%) in the bone marrow, but we deemed these as germline variants given they clustered with the germline variants. The CD3 purity for these three cases was above 89%.

Where somatic variants were detected in the T cells, this was largely explained by contamination of the T cells by other blood cells (Figure S1 and Table S2).

Somatic pathogenic/likely pathogenic variants

Somatic P/LP variants were present in 118 patients: 29% (47/160) of IC patients, who are referred to as Clonal Cytopenia (CC) patients, and 85% (69/81) patients with a myeloid neoplasm (Figure 2; Table S3). The 113 IC patients with no somatic P/LP variants detected were termed Non-Clonal Idiopathic Cytopenia (NCIC) cases. There were significantly more variants per patient in the MN cohort than in the CC cohort (Median 3 variants in MN and 2 in CC, Mann–Whitney test $p = 0.012$). Both CC and MN patients had the same median VAF (36%, Mann–Whitney test $p = 0.44$).

Somatic P/LP variants were discovered in only 39 of the 82 sequenced genes; 17 genes were mutated in 29% of IC cases (47/160), and 37 genes were mutated in 85% (69/81) of MN cases (Figure 3).

Use of T cells as a control aids identification of germline variants

To determine the impact of using a paired reference T-cell sample, we independently analysed 257 patient marrow samples, either individually or paired with T cells (IC $n = 160$, MN $n = 81$ and lymphoid/other $n = 16$), and compared the results.

Reporting guidelines state that variants in the 40%–60% VAF range may be germline or somatic and typically require additional testing to resolve this ambiguity.[21,24] The use of a T-cell control should allow these to be distinguished.[16]

Using the T-cell control, 142 germline variants (10 P/LP and 132 VUS) were discovered in 104 patients. Nine germline *DDX41* P/LP variants (seven in MN cases and two in NCIC cases) and one germline *RUNX1* P/LP variant were identified (Table S4).

Somatic P/LP variants are common in the 40%–60% VAF range

The paired approach was useful to unambiguously identify the origin of the variant. Sixty-six out of 257 sequenced patients had at least one P/LP variant in the 40%–60% VAF range. Paired analysis showed that 56 of these patients exclusively had somatic P/LP variants CC $n = 22/47$ (47%), MN $n = 33/81$ (41%) and lymphoid/other $n = 1/16$ (6%) and 10 had germline P/LP variants.

Fifty-eight out of 257 sequenced patients (25/47 (53%) CC and 32/81 (40%) MN and 1/16 (6%) lymphoid/other patients) had P/LP variants with VAFs, all either <40% or >60%, and paired analysis confirmed their somatic status.

If only the unpaired data were available, the 18 patients harbouring P/LP variants in the 40%–60% VAF range in three genes that are of clinical significance if germline (*DDX41*, *RUNX1 and TP53)* would require additional testing (e.g. skin fibroblast testing) to confirm somatic/germline status.

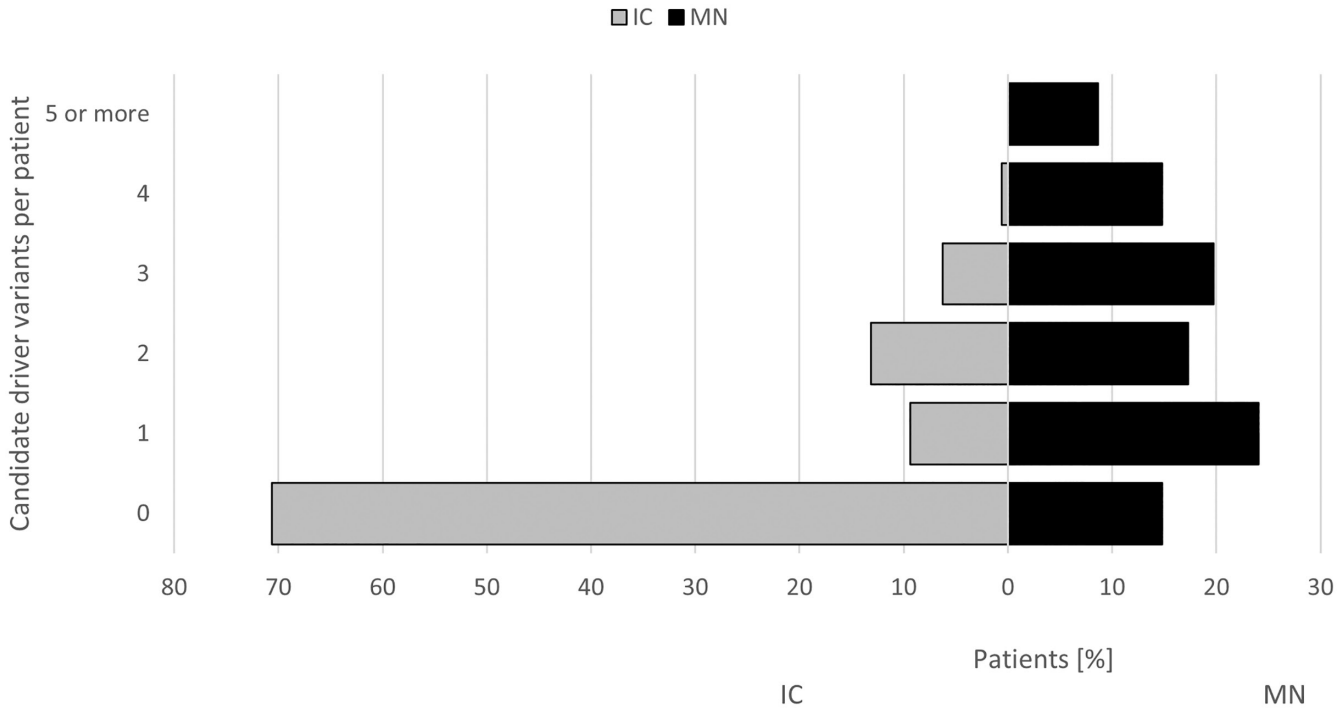

FIGURE 2 Somatic P/LP mutation distribution in IC and MN patients. A histogram of the total number of discovered somatic pathogenic mutations per IC (grey) and MN (black) patient.

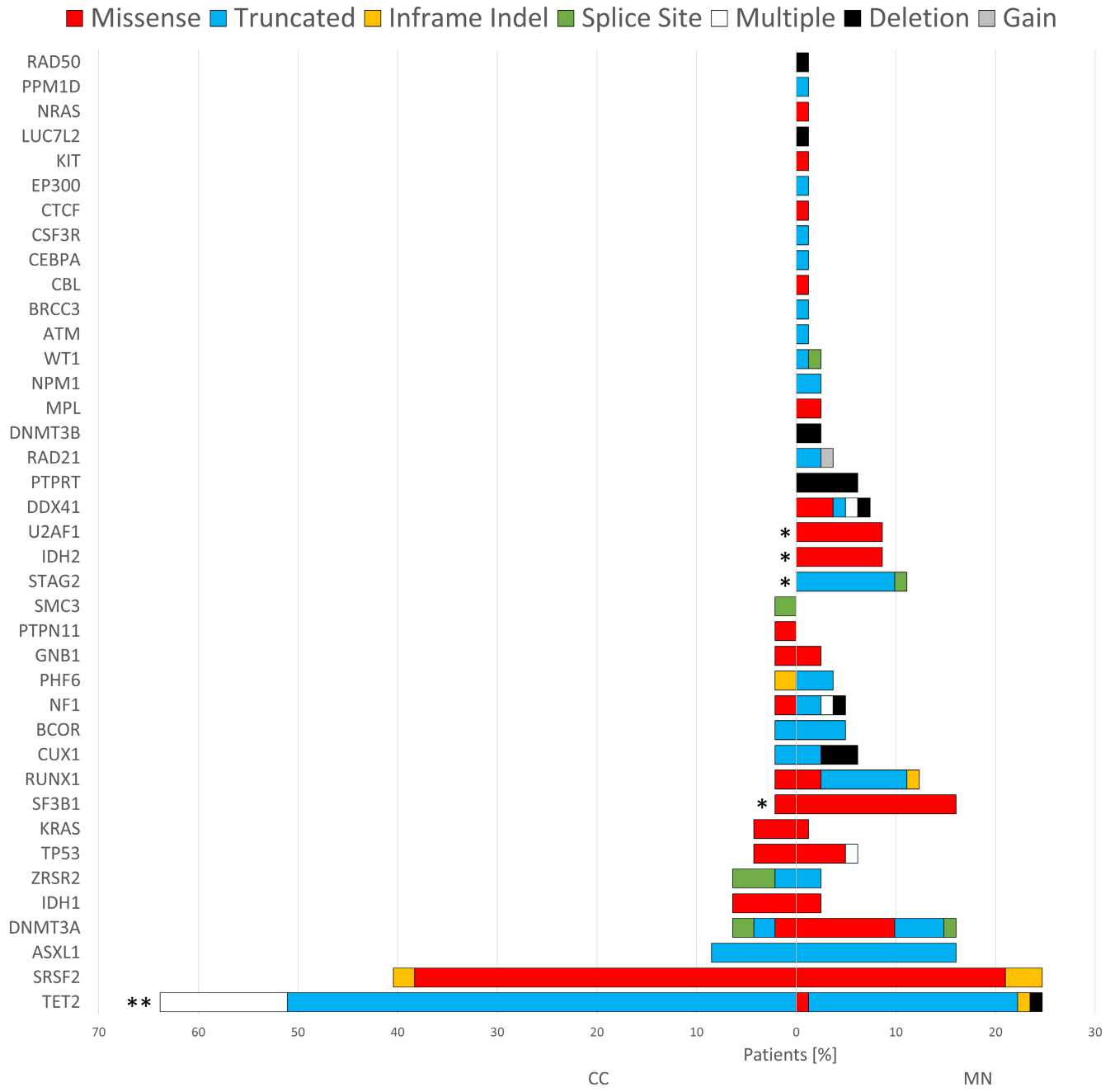

FIGURE 3 Somatic P/LP variant distribution in CC and MN patients. The frequency of P/LP somatic variants in CC patients (left) and MN patients (right) is shown. Colours represent the type of mutation: missense (red), truncated (light blue), inframe deletion (yellow), splice site variant (green), deletion (black), gain (grey) and multiple different categories of variants in the same gene (white). Statistical analysis comparing the frequency of P/LP variants for a given gene between CC and MN patients was performed using Fisher's exact test. * indicates $p < 0.05$, ** indicates $p < 0.0001$.

Somatic VUS are common in clonal myeloid disorders

We discovered VUS in 118 out of 257 sequenced patients. Thirty of the 118 patients had a somatic VUS (IC and MN cases had a median VAF of 31% and 36%, respectively, Mann–Whitney test $p = 0.78$; Tables S5 and S6). Most of the patients (30/33) with a somatic VUS also had a somatic P/LP variant (10/12 IC cases and 20/21 MN cases); only 3 of 33 patients had a somatic VUS with no somatic P/LP variant (2/12 IC patients and 1/21 MN). Somatic VUS were significantly more common in both MN and CC patients than NCIC patients (Somatic VUS were seen in 21% of CC and 26% of MN patients, but only in 1.8% of NCIC patients, Chi square test $p < 0.0001$), consistent with the notion that somatic VUS are indicators of a clonal process. The VUS in the two NCIC patients occurred in myeloid driver genes as defined by the WHO 2022, so hit criteria for CCUS. Ninety-seven of the 118 patients had a germline VUS consistent with benign private variants.

Paired analysis is faster than unpaired

We timed the reporting of some cases to assess the impact of paired analysis on reporting time. The use of a paired reference sample reduced the time required for analysis from median 33 min (IQR: 22–46 min, $n = 16$) for unpaired samples to median 13 min (IQR: 10–24 min, $n = 25$) for paired samples (Mann–Whitney test $p = 0.0005$). Analysis of paired samples takes less time, because it is quicker to identify sequencing errors and artefacts, and distinguish somatic and germline variants.

CD3-depleted peripheral blood (DPB) as an alternative to bone marrow for NGS genetic testing

We tested if DPB generated during T-cell harvesting could be used in place of bone marrow as a source of diagnostic material. DNA isolated from DPB was available for 207 patients (121 IC, 73 MN and 13 lymphoid cases), and was sequenced and compared with the BM result. The VAF of somatic variants was not significantly different in the DPB and BM, with means of 37.1% (±16%) and 37.9% (±17.9%) respectively (paired t-test $p = 0.14$) (Figure 4).

A false negative result was seen in six DPB samples of 106 cases where somatic P/LP variants were detected in the BM, giving a false-negative rate for DPB of 5.8%. Conversely, somatic P/LP variants were detected in three DPB samples that were not detected in the BM. Full concordance, where the same variants were detected in the BM and DPB samples, was seen in 90% of patients, and 96% of patients had partially concordant results (Figure 5).

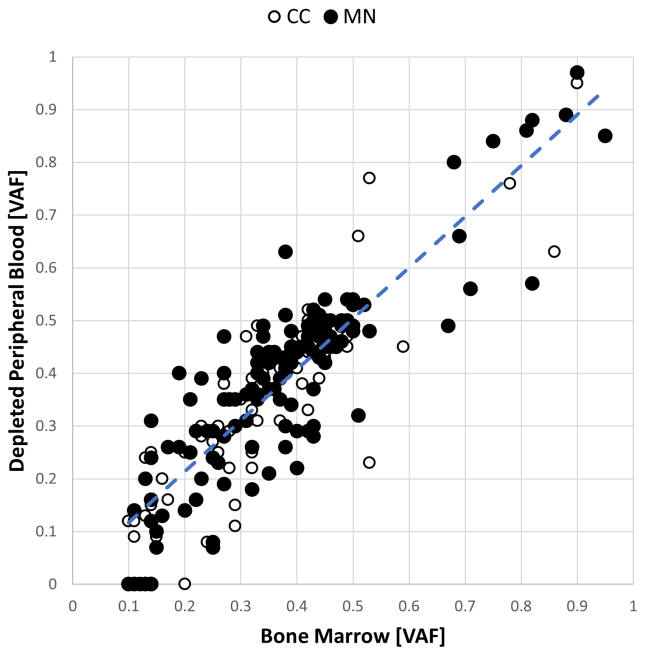

FIGURE 4 Comparison of allele frequencies of somatic mutations discovered in bone marrow and CD3-depleted blood. Black circles represent variants from MN (myeloid neoplasm) cases. White circles represent variants from CC (clonal cytopenia) cases (Pearson coefficient of determination $r^2 = 0.76$; Slope = 0.79; $p < 0.001$). There was no significant difference in mean VAFs between DPB and BM, with means of 37.1% (±16%) and 37.9% (±17.9%) respectively (paired t-test $p = 0.14$).

Impact of lowering VAF threshold to 5% from 10%

We wanted to see the impact of lowering the VAF threshold to 5%, which is within the limits of detection validated for our sequencing approach.

Use of a 5% VAF threshold allowed the discovery of additional somatic P/LP variants in the BM of 33 patients (IC $n = 10/160$, MN $n = 19/81$ and lymphoid/other $n = 4/16$) out of a total of 257 patients. Fifteen out of these 33 patients had variants exclusively at 5%–10% VAF (IC $n = 4/160$, MN $n = 7/81$ and lymphoid/other $n = 4/16$), and 18 patients had other P/LP variants above 10% VAF (Figure 6). One hundred and eighteen patients had P/LP variants above a 10% VAF threshold, compared to 133 patients above the 5% VAF threshold, giving a false-negative rate of 11% (15/133).

Using the 5% VAF threshold affects the classification and prognostication of the patients: four IC patients who would have been classified as non-clonal based on a 10% VAF threshold had P/LP variants at 5%–10% VAF. Seven MDS patients had P/LP variants in IPSS-M prognostic genes with a 5%–10% VAF. The IPSS-M risk category worsened in three of these seven patients when the 5%–10% VAF range P/LP variants were included in the risk score, compared with when only P/LP variants above 10% were included.

We also looked at the 2% VAF threshold in line with the WHO/ICC CCUS definitions and IPSS-M. This moved 7% of NCIC patients with somatic P/LP variants at 2%–5% VAF into the CC category (Figure S4). Thirteen out of 49 MDS patients had additional somatic P/LP variants in the 2%–5% range that score on the IPSS-M.

DISCUSSION

In this study, we used paired tumour-normal analysis to identify somatic P/LP variants in 29% of IC patients. This is comparable to previous studies, which reported that 20%–35% of IC patients harboured mutations using VAF cut-offs between 2% and 10%.[25–28] Here, lowering the VAF threshold to 5% increased the number of IC patients identified as having clonal haemopoiesis. While identifying low-level subclones may be prognostically important in AML/MDS, there is a potential risk if very low VAF thresholds are used while investigating IC patients; cytopenias may be misattributed to low-level clones, which are less likely to be responsible for the cytopenias or have a lower prognostic value.[17,19,29–31]

Factors other than the VAF of variants are prognostic in the context of CCUS. The type of mutation as well as the

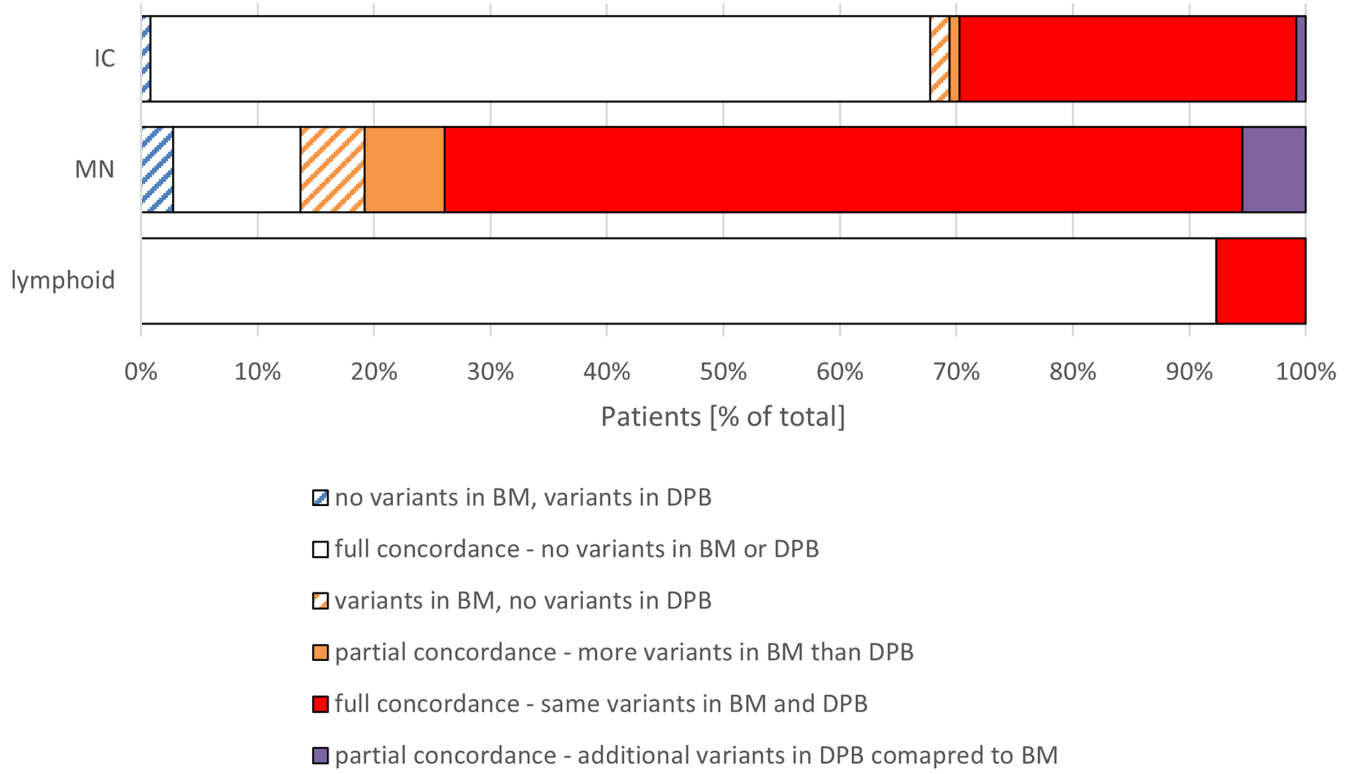

FIGURE 5 Comparison of bone marrow and depleted blood sequencing results at 10% VAF threshold. The bars represent the percentage of analysed patients from each group (IC $n = 121$, MN $n = 73$ and lymphoid $n = 13$). Bars with blue diagonal stripes represent patients who had no reportable variants in BM but had variants in DPB. White bars represent patients who had no variants in both BM and DPB. Bars with orange diagonal stripes represent patients who had variants detected in BM but none in DPB. The orange bar represents partially concordant patients who had some, but not all, variants present in BM detected in DPB. The red bar represents fully concordant patients who had the same variants detected in BM and DPB. The purple bar represents partially concordant patients who had the same variants present in BM detected in DPB, with additional variants detected in DPB.

number of mutations can be used to classify CCUS into high- and low-risk categories.[28,31]

We discovered P/LP variants in only 39 of the 82 genes tested, similar to one study that reported that 41 genes provided similar diagnostic information to 640 genes.[27] Clonal cytopenias are defined by the presence of somatic VUS as well as somatic P/LP variants in 54 myeloid driver genes listed in the WHO 2022; therefore, maintaining a broader panel of genes may be of utility.

We demonstrated largely concordant results between bone marrow and DPB, providing further evidence for routine molecular profiling of blood. Although previous studies have compared sequencing of peripheral blood and bone marrow, this is the first study that used material acquired at the same timepoint, and employed a capture-based panel in combination with lymphocyte depletion to enrich for myeloid cells (to reduce the risk of a false negative where patients have profound neutropenia/monocytopenia).[32–34] Although the investigation of cytopenias could start with the sequencing of DPB, there is a danger of missing some cases of MDS as not all bear mutations are detectable by current approaches.[28]

The main benefits of the paired approach were both reliable identification and interpretation of germline variants and a reduction of analysis time, as the assessment of VUS and the identification of artefacts, misalignments and errors were greatly simplified. Timely identification of germline variants predisposing to haematological malignancies is of particular relevance for the selection of family donors for transplant, as donor-derived leukaemia can arise in recipients where donors have germline pathogenic variants.[35] Sequencing fibroblasts cultured from skin biopsies is an alternative approach. Use of T cells is quicker and avoids the need for skin biopsy and culture of the cells, but it does involve immunomagnetic selection, and selection of T cells from some AML samples can be difficult.[36] The paired approach increased costs by 85% compared to unpaired sequencing only, but was less expensive than fibroblast culture, which would increase costs by 160%. Units could employ a strategy of collecting T-cell DNA but only sequencing it if ambiguous variants in potentially germline variants were identified in the marrow. Patients with family donors who are moving to transplant as first treatment or for those who go to transplant after one remission induction chemotherapy cycle may, however, experience delays if the T-cell sequencing is not performed simultaneously with the marrow.

We confirmed the DDX41 germline variant in one of the 10 germline cases through skin fibroblast testing. We cannot fully exclude the possibility that some of the other

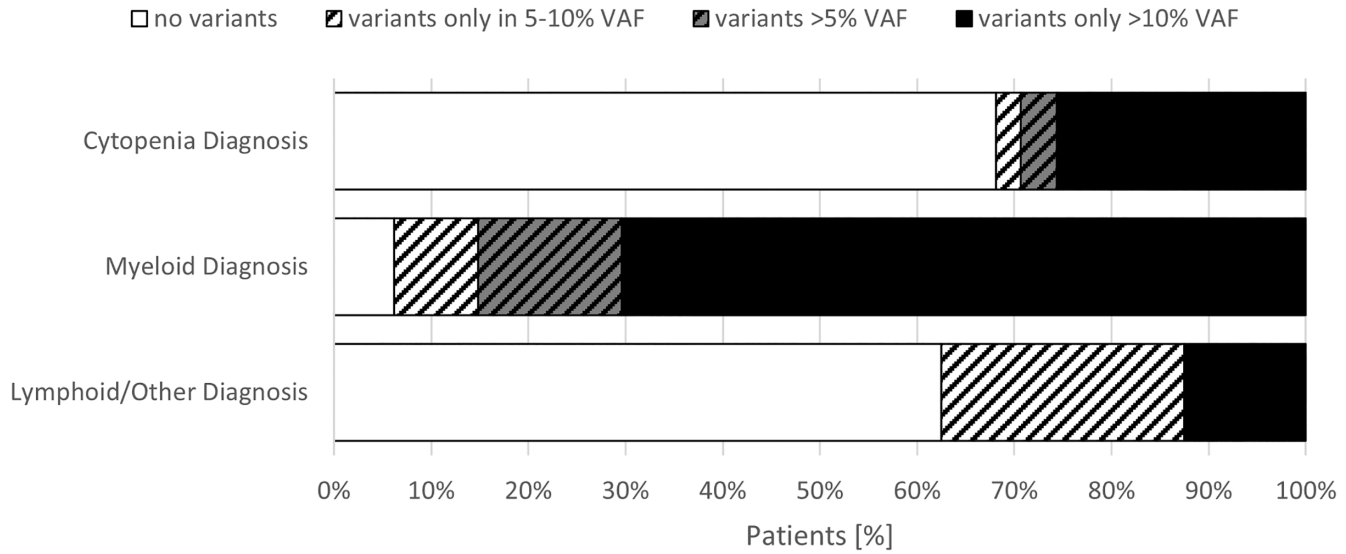

FIGURE 6 Comparison between 5% and 10% VAF reporting threshold. The bars represent the percentage of analysed patients from each group (IC $n = 160$, MN $n = 81$ and lymphoid/other $n = 16$). Patients were divided into four categories: no P/LP variants (white), P/LP variants only in the 5%–10% VAF range (white with black diagonal stripes), P/LP variants in both 5%–10% and >10% ranges (grey with black diagonal stripes) and patients with P/LP variants only above 10% VAF (black).

variants we deemed to be germline occurred somatically at the haematopoietic stem cell (HSC) level and hence were detected in T cells and bone marrow. Against this, there is clear space in terms of the VAF between the somatic variants in the T cells, with a maximum VAF of 22%, and the germline variants, with a minimum VAF of 38%. Other studies show low VAF (<25%) of somatic variants in the T cells where HSC involvement was suspected.[37–39] We are unaware of any studies that show high VAF (>35%) involvement in T cells as a result of HSC involvement by somatic myeloid drivers.

We observed somatic P/LP variants in *TET2* and *SRSF2* in CC patients at a higher frequency than in other studies (64% vs. 15%–30% for *TET2* and 40% vs. 0%–20% for *SRSF2*).[25–27,40,41] This discrepancy could be due to differences in the classification of ambiguous variants.[42] A study comparing variant interpretations between laboratories found that 11% of patients had the same variants classified differently, ranging from pathogenic to VUS.[43] An example is the *SRSF2* hotspot p.Pro95, where only one out of the three somatic variants (p.Pro95Arg) is consistently classified as P/LP while the other two (p.Pro95His and p.Pro96Leu) are classified differently between databases, from likely pathogenic to VUS.[1,44] Such ambiguity could explain the different frequencies of mutated genes reported between studies, as usually only P/LP variants are reported, and VUS is omitted from analysis.

We detected somatic VUS in a significantly higher proportion of patients with MN and CC than patients with NCIC, consistent with somatic VUS being an indicator of a clonal process. The importance of VUS is a hotly debated topic, and significant effort is placed towards resolving this ambiguity.[45,46] The transparent reporting of VUS may help resolve their significance.[42]

AUTHOR CONTRIBUTIONS

Study design: DCT, GP, NC, MH, KM, SOC, DW and JK; Study Investigator: DCT, GP, CDL, GM, BC, SA, MF, JOC, EGD, SN, AMcG, MK, JB, SI, MMcM, TC, AP, CJ, JB and SOC; Contributed patients or study materials: DCT, CDL, GM, BC, SA, MF, JOC, EGD, SN, AmcG, MK, JB, MMcM, TC, AP, CJ and JB; Collection and assembly of data: DCT and GP; Analysed the data: GP, JK, DW, SR, JK, KM and NC; Data interpretation: DCT, GP, JK, SR, DW, KM and NC; Critical review and revision of this manuscript and approval of the manuscript for submission: All authors.

AFFILIATIONS

[1]Division of Molecular Pathology, Institute of Cancer Research, London, UK
[2]Department of Translational Research, Royal Marsden Hospital NHS Trust, London, UK
[3]Department of Haematology, St Helier Hospital, London, UK
[4]Department of Haematology, Royal Marsden Hospital NHS Trust, London, UK
[5]Department of Haematology, Queen Alexandra Hospital, Portsmouth, UK
[6]Department of Haematology, Croydon University Hospital, London, UK
[7]Department of Haematology, Kingston Hospital NHS Foundation Trust, London, UK
[8]Department of Haematology, Royal Cornwall Hospitals NHS Foundation Trust, Cornwall, Truro, UK
[9]Department of Haematology, Salisbury NHS Foundation Trust, Salisbury, UK
[10]Department of Haematology, Royal Surrey County Hospital NHS Foundation Trust, Guildford, UK
[11]Department of Haematology, University Hospital Southampton, Southampton, UK
[12]Department of Haematology, Newcastle Upon Tyne Hospitals NHS Foundation Trust, Newcastle Upon Tyne, UK
[13]Department of Haematology, Northumbria Healthcare NHS Foundation Trust, Tyneside, UK
[14]Department of Haematology, Queen's University, Belfast, UK
[15]Department of Haematology, Royal Devon and Exeter NHS Foundation Trust, Exeter, UK
[16]Department of Haematology, Northwick Park Hospital, London, UK
[17]Department of Haematology, St Bartholomew's Hospital, London, UK
[18]Department of Haematology, Nottingham University Hospitals NHS Trust, Nottingham, UK
[19]Department of Statistics, Royal Marsden Hospital NHS Trust, London, UK

[20]Faculty of Medicine, University of Southampton, Southampton, UK
[21]Wessex Genomics Laboratory Service, Salisbury NHS Foundation Trust, Salisbury, UK
[22]Department of Translational Research, Royal Marsden Hospital Sutton, London, UK
[23]Division of Molecular Pathology, Clinical Genomics (Research), Institute of Cancer Research, London, UK
[24]Department of Molecular Pathology, Royal Marsden Hospital Sutton, London, UK
[25]Department of Molecular Pathology, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK
[26]Department of Histopathology, Royal Marsden NHS Foundation Trust, London, UK

ACKNOWLEDGEMENTS

The project was funded by grant A127 through the National Institute for Health Research at the Royal Marsden/Institute of Cancer Research Biomedical Research Centre and a grant from the Royal Marsden Charity.

FUNDING INFORMATION

The project was funded by grant A127 through the National Institute for Health Research at the Royal Marsden/Institute of Cancer Research Biomedical Research Centre and a grant from the Royal Marsden Charity.

CONFLICT OF INTEREST STATEMENT

None of the authors has a relevant conflict of interest.

DATA AVAILABILITY STATEMENT

N/A.

ETHICS STATEMENT

The 'Improving diagnosis in idiopathic cytopenia using gene sequencing' protocol was reviewed by the South Yorkshire Research Ethics Committee (REC reference 16/YH/1042, IRAS number 181550). The trial sponsor was the RMH (CCR4384). Additional work was covered by the 'Laboratory Studies into the pathology of leukaemia' protocol (REC reference 16/EE/0266, IRAS number 197096) reviewed by the East of England–Cambridge South Research Ethics Committee.

PATIENT CONSENT STATEMENT

The study was performed according to the Helsinki Declaration. Patients provided written informed consent.

PERMISSION TO REPRODUCE MATERIAL FROM OTHER SOURCES

N/A.

CLINICAL TRIAL REGISTRATION (INCLUDING TRIAL NUMBER)

NCT03026751.

ORCID

Mary Frances McMullin https://orcid.org/0000-0002-0773-0204
Thomas Coats https://orcid.org/0000-0003-4341-4361
Nicholas Cross https://orcid.org/0000-0001-5481-2555
David Taussig https://orcid.org/0000-0002-9621-0146

REFERENCES

1. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res. 2018;46(D1):D1062–D1067.
2. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, et al. COSMIC: the catalogue of somatic mutations in cancer. Nucleic Acids Res. 2019;47(D1):D941–D947.
3. Chakravarty D, Gao J, Phillips S, Kundra R, Zhang H, Wang J, et al. OncoKB: a precision oncology knowledge base. JCO Precis Oncol. 2017;16(1):1–9.
4. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141 456 humans. Nature. 2020;581(7809):434–43.
5. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. 2001;29(1):308–11.
6. Teer JK, Zhang Y, Chen L, Welsh EA, Cress WD, Eschrich SA, et al. Evaluating somatic tumour mutation detection without matched normal samples. Hum Genomics. 2017;11:22.
7. Jones S, Anagnostou V, Lytle K, Parpart-Li S, Nesselbush M, Riley DR, et al. Personalized genomic analyses for cancer mutation discovery and interpretation. Sci Transl Med. 2015;7(283):283ra53.
8. McKerrell T, Moreno T, Ponstingl H, Bolli N, Dias JML, Tischler G, et al. Development and validation of a comprehensive genomic diagnostic tool for myeloid malignancies. Blood. 2016;128(1):e1–e9.
9. Haferlach T, Nagata Y, Grossmann V, Okuno Y, Bacher U, Nagae G, et al. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. Leukaemia. 2014;28(2):241–7.
10. Papaemmanuil E, Gerstung M, Malcovati L, Tauro S, Gundem G, Van Loo P, et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood. 2013;122(22):3616–27.
11. Bacher U, Haferlach T, Kern W, Haferlach C, Schnittger S. A comparative study of molecular mutations in 381 patients with myelodysplastic syndrome and in 4130 patients with acute myeloid leukaemia. Haematologica. 2007;92(6):744–52.
12. Lin PH, Li HY, Fan SC, Yuan TH, Chen M, Hsu YH, et al. A targeted next-generation sequencing in the molecular risk stratification of adult acute myeloid leukaemia: implications for clinical practice. Cancer Med. 2017;6:349–60.
13. Thol F, Kade S, Schlarmann C, Löffeld P, Morgan M, Krauter J, et al. Frequency and prognostic impact of mutations in SRSF2, U2AF1, and ZRSR2 in patients with myelodysplastic syndromes. Blood. 2012;119(15):3578–84.
14. Abdel-Wahab O, Figueroa ME. Interpreting new molecular genetics in myelodysplastic syndromes. ASH Educ Program Book. 2012;2012(1):56–64.
15. Bejar R, Stevenson K, Abdel-Wahab O, Galili N, Nilsson B, Garcia-Manero G, et al. Clinical effect of point mutations in myelodysplastic syndromes. N Engl J Med. 2011;364(26):2496–506.
16. Score J, Chase A, Forsberg LA, Feng L, Waghorn K, Jones AV, et al. Detection of leukaemia-associated mutations in peripheral blood DNA of hematologically normal elderly individuals. Leukaemia. 2015;29(7):1600–2.
17. Young AL, Challen GA, Birmann BM, Druley TE. Clonal haematopoiesis harbouring AML-associated mutations is ubiquitous in healthy adults. Nat Commun. 2016;7(1):12484.
18. Miraki-Moud F, Anjos-Afonso F, Hodby KA, Griessinger E, Rosignoli G, Lillington D, et al. Acute myeloid leukaemia does not deplete normal haematopoietic stem cells but induces cytopenias by impeding their differentiation. Proc Natl Acad Sci. 2013;110(33):13576–81.
19. Buscarlet M, Provost S, Zada YF, Barhdadi A, Bourgoin V, Lépine G, et al. DNMT3A and TET2 dominate clonal hematopoiesis and demonstrate benign phenotypes and different genetic predispositions. Blood. 2017;130(6):753–62.
20. Rabizadeh S, Garner C, Sanborn JZ, Benz SC, Reddy S, Soon-Shiong P. Comprehensive genomic transcriptomic tumour-normal gene

panel analysis for enhanced precision in patients with lung cancer. Oncotarget. 2018;9(27):19223–32.
21. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17(5):405–23.
22. Jarvik GP, Browning BL. Consideration of Cosegregation in the pathogenicity classification of genomic variants. Am J Hum Genet. 2016;98(6):1077–81.
23. Ellard S, Baple E, Berry I, Forrester N, Turnbull C, Owens M, et al. ACGS best practice guidelines for variant classification 2019. In 2019 [cited 2023 Apr 27]. Available from: https://www.semanticscholar.org/paper/ACGS-Best-Practice-Guidelines-for-Variant-2019-Ellard-Baple/86fa75f316823ef358a804fe1e45a98a85351bf9
24. Li MM, Datto M, Duncavage EJ, Kulkarni S, Lindeman NI, Roy S, et al. Standards and guidelines for the interpretation and reporting of sequence variants in cancer: a joint consensus recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. J Mol Diagn JMD. 2017;19(1):4–23.
25. Malcovati L, Gallì A, Travaglino E, Ambaglio I, Rizzo E, Molteni E, et al. Clinical significance of somatic mutation in unexplained blood cytopenia. Blood. 2017;129(25):3371–8.
26. Kwok B, Hall JM, Witte JS, Xu Y, Reddy P, Lin K, et al. MDS-associated somatic mutations and clonal hematopoiesis are common in idiopathic cytopenias of undetermined significance. Blood. 2015;126(21):2355–61.
27. Zheng G, Chen P, Pallavajjalla A, Haley L, Gondek L, Dezern A, et al. The diagnostic utility of targeted gene panel sequencing in discriminating etiologies of cytopenia. Am J Hematol. 2019;94(10):1141–8.
28. Cargo C, Bernard E, Beinortas T, Bolton KL, Glover P, Warren H, et al. Predicting cytopenias, progression, and survival in patients with clonal cytopenia of undetermined significance: a prospective cohort study. Lancet Haematol. 2024;11:e51–e61.
29. Gallì A, Todisco G, Catamo E, Sala C, Elena C, Pozzi S, et al. Relationship between clone metrics and clinical outcome in clonal cytopenia. Blood. 2021;138(11):965–76.
30. Rossi M, Meggendorfer M, Zampini M, Tettamanti M, Riva E, Travaglino E, et al. Clinical relevance of clonal hematopoiesis in persons aged ≥80 years. Blood. 2021;138(21):2093–105.
31. Weeks LD, Niroula A, Neuberg D, Wong W, Lindsley RC, Luskin MR, et al. Prediction of risk for myeloid malignancy in clonal hematopoiesis. NEJM Evid. 2023;2(5):EVIDoa2200310.
32. Mohamedali AM, Gäken J, Ahmed M, Malik F, Smith AE, Best S, et al. High concordance of genomic and cytogenetic aberrations between peripheral blood and bone marrow in myelodysplastic syndrome (MDS). Leukaemia. 2015;29(9):1928–38.
33. Jumniensuk C, Nobori A, Lee T, Senaratne TN, Rao D, Pullarkat S. Concordance of peripheral blood and bone marrow next-generation sequencing in Haematological neoplasms. Adv Hematol. 2022;2022:8091746.
34. Hogg G, Severson EA, Cai L, Hoffmann HM, Holden KA, Fitzgerald K, et al. Clinical characterization of the mutational landscape of 24 639 real-world samples from patients with myeloid malignancies. Cancer Genet. 2023;278:38–49.
35. Xiao H, Shi J, Luo Y, Tan Y, He J, Xie W, et al. First report of multiple CEBPA mutations contributing to donor origin of leukaemia relapse after allogeneic haematopoietic stem cell transplantation. Blood. 2011;117(19):5257–60.
36. Le Dieu R, Taussig D, Lister TA, Gribben JG. Negative immunomagnetic selection of T cells from peripheral blood of presentation AML specimens. J Immunol Methods. 2009;348(1):95–100.
37. Shlush LI, Zandi S, Mitchell A, Chen WC, Brandwein JM, Gupta V, et al. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. Nature. 2014;506(7488):328–33.
38. Beck DB, Ferrada MA, Sikora KA, Ombrello AK, Collins JC, Pei W, et al. Somatic mutations in UBA1 and severe adult-onset autoinflammatory disease. N Engl J Med. 2020;383(27):2628–38.
39. Corces-Zimmerman MR, Hong WJ, Weissman IL, Medeiros BC, Majeti R. Preleukemic mutations in human acute myeloid leukaemia affect epigenetic regulators and persist in remission. Proc Natl Acad Sci U S A. 2014;111(7):2548–53.
40. Choi EJ, Cho YU, Hur EH, Park HS, Choi Y, Lee JH, et al. Clinical implications and genetic features of clonal cytopenia of undetermined significance compared to lower-risk myelodysplastic syndrome. Br J Haematol. 2022;198(4):703–12.
41. Li M, Binder M, Lasho T, Ferrer A, Gangat N, Al-Kali A, et al. Clinical, molecular, and prognostic comparisons between CCUS and lower-risk MDS: a study of 187 molecularly annotated patients. Blood Adv. 2021;5(8):2272–8.
42. Moody EW, Vagher J, Espinel W, Goldgar D, Hagerty KJ, Gammon A. Comparison of somatic and germline variant interpretation in hereditary cancer genes. JCO Precis Oncol. 2019;3:1–8.
43. Balmaña J, Digiovanni L, Gaddam P, Walsh MF, Joseph V, Stadler ZK, et al. Conflicting interpretation of genetic variants and cancer risk by commercial laboratories as assessed by the prospective registry of multiplex testing. J Clin Oncol. 2016;34(34):4071–8.
44. Kopanos C, Tsiolkas V, Kouris A, Chapple CE, Albarca Aguilera M, Meyer R, et al. VarSome: the human genomic variant search engine. Bioinformatics. 2019;35(11):1978–80.
45. Federici G, Soddu S. Variants of uncertain significance in the era of high-throughput genome sequencing: a lesson from breast and ovary cancers. J Exp Clin Cancer Res CR. 2020;39:46.
46. Horak P, Griffith M, Danos AM, Pitel BA, Madhavan S, Liu X, et al. Standards for the classification of pathogenicity of somatic variants in cancer (oncogenicity): joint recommendations of clinical genome resource (ClinGen), cancer genomics consortium (CGC), and variant interpretation for cancer consortium (VICC). Genet Med off J Am Coll Med Genet. 2022;24(5):986–98.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Pietka G, De Lord C, Matthias G, Cheung B, Atwal S, Furtado M, et al. Capture-based targeted sequencing using a T-cell control in myeloid malignancies and idiopathic cytopenias. Br J Haematol. 2024;00:1–10. https://doi.org/10.1111/bjh.19377