The spectrum and clinical impact of epigenetic modifier mutations in myeloma

Pawlyn, Charlotte; Kaiser, Martin F; Heuck, Christoph; Melchor, Lorenzo; Wardell, Christopher P; Murison, Alex; Chavan, Shweta S; Johnson, David C; Begum, Dil B; Dahir, Nasrin M; Proszek, Paula Z; Cairns, David A; Boyle, Eileen M; Jones, John R; Cook, Gordon; Owen, Roger G; Gregory, Walter M; Jackson, Graham H; Barlogie, Bart; Davies, Faith E

DOI:
10.1158/1078-0432.CCR-15-1790

The spectrum and clinical impact of epigenetic modifier mutations in myeloma

Running title: Epigenetic modifier mutations in myeloma

Authors

Charlotte Pawlyn[1], Martin F Kaiser[1], Christoph Heuck[2], Lorenzo Melchor[1], Christopher P Wardell[1*], Alex Murison[1], Shweta Chavan[2], David C Johnson[1], Dil Begum[1], Nasrin Dahir[1], Paula Proszek[1†], David A Cairns[3], Eileen M Boyle[1], John R Jones[1], Gordon Cook[4], Mark T Drayson[5], Roger G Owen[6], Walter M Gregory[3], Graham H Jackson[7], Bart Barlogie[2], Faith E Davies[1,2], Brian A Walker[1,2], Gareth J Morgan[1,2]

Affiliations

[1]The Institute of Cancer Research, London, United Kingdom

[2]Myeloma Institute, University of Arkansas for Medical Sciences, Little Rock, AR, United States

[3]Clinical Trials Research Unit, Leeds Institute of Clinical Trials Research, University of Leeds, Leeds, United Kingdom

[4]University of Leeds, Leeds, United Kingdom

[5]Clinical Immunology, School of Immunity & Infection, University of Birmingham, Birmingham, United Kingdom

[6]St James's University Hospital, Leeds, United Kingdom

[7] Department of Haematology, Newcastle University, Newcastle, United Kingdom

†Current address: Centre for Molecular Pathology, The Royal Marsden NHS Foundation Trust, London, United Kingdom

*Current address: Laboratory for Genome Sequencing Analysis, RIKEN Center for Integrative Medical Sciences, Tokyo, Japan

Key words:

myeloma, whole-exome sequencing, epigenetic modifiers

Corresponding author:

Professor Gareth J Morgan

- The Institute of Cancer Research, 15, Cotswold Rd, London, SM2 5NG. United Kingdom

  Telephone: +44 2087224130 Fax: +44 2087224432

- Myeloma Institute for Research and Therapy, University of Arkansas for Medical Sciences, 4301 West Markham, #816, Little Rock, Arkansas 72205. United States

  E-mail: GJMorgan@uams.edu

  Telephone: +1 501 526 6990 ext. 2456 Fax: +1 501 526 2273

Conflict of Interest Statement

The authors declare no relevant conflict of interest.

Translational relevance

Myeloma is a heterogeneous malignancy with disease in different subgroups of patients driven by abnormalities in multiple genes and/or molecular pathways. Treatment options and outcomes have improved over the last decade but novel approaches are still required. In this article we use whole-exome sequencing results from 463 presenting cases entered into the UK Myeloma XI study, and targeted sequencing of 156 previously treated cases, to report the wide spectrum of mutations in genes encoding epigenetic modifiers in myeloma. Using linked survival data from the large Myeloma XI study we identify lesions that may have prognostic significance in *KDM6A* and genes encoding DNA modifiers. We demonstrate an increase in the frequency of epigenetic modifier mutations of certain classes as disease progresses. Our analysis is particularly important as numerous mutations identified suggest potential targeted treatment strategies with agents either currently available or known to be in development, highlighting novel treatment approaches for patients.

Abstract

Purpose: Epigenetic dysregulation is known to be an important contributor to myeloma pathogenesis but, unlike in other B cell malignancies, the full spectrum of somatic mutations in epigenetic modifiers has not been previously reported. We sought to address this using results from whole-exome sequencing in the context of a large prospective clinical trial of newly diagnosed patients and targeted sequencing in a cohort of previously treated patients for comparison.

Experimental Design: Whole-exome sequencing analysis of 463 presenting myeloma cases entered in the UK NCRI Myeloma XI study and targeted sequencing analysis of 156 previously treated cases from the University of Arkansas for Medical Sciences. We correlated the presence of mutations with clinical outcome from diagnosis and compared the mutations found at diagnosis with later stages of disease.

Results: In diagnostic myeloma patient samples we identify significant mutations in genes encoding the histone 1 linker protein, previously identified in other B-cell malignancies. Our data suggest an adverse prognostic impact from the presence of lesions in genes encoding DNA methylation modifiers and the histone demethylase *KDM6A/UTX*. The frequency of mutations in epigenetic modifiers appears to increase following treatment most notably in genes encoding histone methyltransferases and DNA methylation modifiers.

Conclusions: Numerous mutations identified raise the possibility of targeted treatment strategies for patients either at diagnosis or relapse supporting the use of sequencing-based diagnostics in myeloma to help guide therapy as more epigenetic targeted agents become available.

Introduction

Myeloma is a malignancy of plasma cells, terminally differentiated B cells involved in the immune response. Despite advances in therapy over the last 10 years subgroups of patients diagnosed with myeloma continue to have poor outcomes and most inevitably relapse. A better understanding of the genetic and epigenetic abnormalities that contribute to disease pathogenesis is required in order to develop new targeted treatment strategies.

The myeloma clone is thought to be immortalised following the acquisition of a translocation into the immunoglobulin heavy chain locus (t(4;14), t(6;14), t(11;14), t(14;16) and t(14;20)) or hyperdiploidy. The clone then evolves via the subsequent gain of further genetic or epigenetic events.(1) Epigenetic dysregulation is known to be an important contributor to myeloma progression, but the full extent of its role in the pathogenesis of disease and high risk behaviour is uncertain.(2) A key example of the relevance and role of epigenetic deregulation in myeloma comes from our understanding of the subgroup of patients with the t(4;14), which results in the juxtaposition of *MMSET*, a H3K36 methyltransferase, to the immunoglobulin heavy chain locus enhancer. The resulting overexpression of MMSET leads to a histone methylation pattern, characterised by increased H3K36me2 and decreased H3K27me3, along with a distinct and aberrant DNA methylation pattern.(3) Downstream events occurring as a result of this primary event include changes in expression of genes involved in the cell cycle, apoptosis and cell adhesion.(4-7) Patients carrying the t(4;14), around 15% of myeloma cases at clinical diagnosis(8, 9), tend to respond to treatment but relapse early and have a shorter overall survival,

though since the use of proteasome inhibition there is evidence that this poor prognosis is ameliorated.

The impact of mutation on epigenetic modifiers in myeloma has also been reported, with inactivation of the histone demethylase, *KDM6A/UTX,* seen in 10% of samples (10), and mutations affecting the histone methyltransferase, *MLL*(11) the best documented. The recognition of mutations in *MLL* followed the first sequencing study of myeloma patients where expression of the homeobox protein, HOX9A, was suggested to be regulated by *MLL.(11)* More recently an activating mutation of *MMSET* at E1099K was found in the MM.1S myeloma cell line with the same downstream effect as *MMSET* overexpression due to the t(4;14). (12)

Based on these considerations we sought to investigate the role of mutations in epigenetic modifiers in myeloma and how such mutations might contribute to disease pathogenesis. We used whole exome sequencing to examine the spectrum of mutations in epigenetic modifiers in a series of 463 newly diagnosed patients uniformly treated as part of the UK NCRI Myeloma XI clinical trial (MyXI) and describe the clinical implication of mutations in terms of their effect on progression-free and overall survival. Previously reported results from this dataset identified 15 significantly mutated genes in myeloma, a distinct APOBEC signature associated with maf translocations and the link between these factors and prognosis.(13, 14) Here we extend these analyses with a focus on highlighting epigenetic mutations important in the pathogenesis of myeloma and compare the frequency of mutations at diagnosis to later stages of disease, using a dataset from the University of Arkansas for Medical Sciences (UAMS). This analysis comprises the largest series of newly diagnosed myeloma patients sequenced to date and provides important

insights into the role of epigenetics in the disease, as well as highlighting potential avenues for future research and targeted treatment development.

Materials and Methods

Whole exome sequencing at diagnosis – Myeloma XI

Methods used for the analysis of patient material have been previously published (13, 14) but are summarised below:

Samples were taken, following informed consent and prior to treatment commencement, from 463 patients newly diagnosed with symptomatic myeloma and enrolled in the UK NCRI Myeloma XI trial (NCT01554852). The study was approved by the NHS Health Research Authority, National Research Ethics Service Committee and by local review committees at all participating centres. The study randomised patients between triplet immunomodulatory drug (IMiD) inductions with thalidomide vs lenalidomide, prior to subsequent randomisations comparing consolidation and maintenance approaches. The demographics of the patients included in this analysis have been published and are reproduced in **Supplementary Table 1**. (14) Progression-free (PFS) and overall survival (OS), measured from initial randomisation, had median follow up of 25 months, 95% CI [24.3,26.2]. The median PFS was 26.6 months, 95%CI [23.6,29.9] and the median OS was not reached but the 3 year OS was 66%, 95%CI [60,73].

CD138+ plasma cells were isolated from bone marrow cells using MACSorting (Miltenyi Biotech, Bisley, UK), lysed in RLT+ buffer and DNA/RNA extracted using the AllPrep kit (Qiagen, Manchester, UK). White blood cells were isolated from

peripheral blood, purified by Ficoll-Pacque and DNA extracted using the QIAamp DNA mini kit (Qiagen).

DNA from both tumour and peripheral blood samples was used in the exome capture protocol.(13, 15) RNA baits were designed against the human exome with additional custom baits tiling the IGH, IGK, IGL and MYC loci to detect the major translocations. Four exome samples were pooled and run on one lane of a HiSeq 2000 (Illumina, Hinxton, UK) using 76-bp paired end reads. Data quality and metrics processing, processes for somatic mutation calling and molecular and copy number assessments are described fully elsewhere(13).

Copy Number Estimation and Cancer Clonal Fraction – Myeloma XI

Copy number across the exome was determined using Control-FREEC(16) utilizing 500bp bins, each overlapping with the subsequent and previous 250bp. A minimum average read depth of 50 was required in the control samples, with at least two neighbouring bins required to show CN aberration to call a region as gained or lost. Copy number profiles for a series of 26 chromosomal regions were compared with copy number values previously estimated by multiplex-ligated probe amplification tests.(17) To ensure a reliable estimation of copy number, only cases with an F correlation above 50% were used for subsequent analyses (370/463).

Cancer clonal fraction was calculated, (18) as the proportion of tumour cells containing an SNV using the equation:

$$CCF = \min(1, \frac{CN.r}{R})$$

Where CCF=cancer clonal fraction (proportion of cells containing the mutation), CN=copy number at that site, r=number of reads containing the mutation at that site, R=total number of reads at that site.

Further data analysis – Myeloma XI

Lists of epigenetic modifiers were curated using database searches and previous publications (**Supplementary Table 2**) with genes divided into eight classes: Core/linker histones (n=79), histone demethylases (n=21), histone methyltransferases (n=40), histone deacetylases (n=20), histone acetyltransferases (n=25), DNA modifiers (n=8), readers (n=17) and chromatin assembly/remodelling (n=46). These lists were used to interrogate the results of the sequencing analysis in order to calculate the percentage of patients with a mutation in each gene and in each class.

For all genes mutated in more than 1% of patients, mutations were mapped to the relevant regions of the protein using Protein Paint (19) and the Catalogue of Somatic Mutations in Cancer (COSMIC) database was searched to look for previously identified variants at the same amino acid residue in other tumours.(20) Multiple sequence alignment of histone 1 genes was performed by inputting sequences from FASTA files obtained from uniprot.org into ClustalW2. Mutations were annotated in GeneDoc. The likely effect of mutations was assessed using SIFT analysis. (21)

Deletions in *KDM6A* were identified using an algorithm to detect deletion of whole exons (windows defined as the regions used in the Agilent exome capture). This was done by comparing the read depth between the tumour and normal samples. The mean depth across the window was required to be > 0.2 of the median depth in the normal sample and < 0.06 in the tumour sample, with the normal value being at

least 8x greater than the tumour value. Positive findings using this method were confirmed or excluded following visualisation in Integrated Genome Viewer (IGV).

Survival curves were plotted (Kaplan-Meier) and the statistical significance of the difference between curves tested using the Logrank test, with $P<0.05$ taken as the level of significance. Multivariate analysis was performed using the cox-regression model inputting the epigenetic mutations with a significant impact on survival by univariate Logrank statistic and other standard clinical factors known to influence survival in myeloma patients.

Targeted sequencing in previously treated patients – UAMS

156 patients who had previously undergone treatment had bone marrow samples taken and genomic profiling (FoundationOne Heme®; Foundation Medicine) performed as part of their disease reassessment work-up. Review of this data was approved by the UAMS institutional review board. The demographics of the patients included in this analysis are shown **Supplementary Table 3**. CD138+ cells were isolated from bone marrow aspirates as previously (22) with DNA and RNA extracted using the Puregene and RNeasy kits (Qiagen) respectively. Samples were submitted to a CLIA-certified, New York State and CAP-accredited laboratory (Foundation Medicine, Cambridge MA) for NGS-based genomic profiling. Hybridization capture was applied to ≥ 50ng of extracted DNA or RNA for 405 cancer related genes and select intronic regions from 31 genes (FoundationOne Heme DNA only, n=405); targeted RNA-seq for rearrangement analysis was performed for 265 genes frequently rearranged in cancer. Sequencing of captured libraries was performed (Illumina HiSeq 2500) to a median exon coverage depth of >250x, and resultant sequences were analyzed for base substitutions, insertions, deletions, copy number

alterations (focal amplifications and homozygous deletions) and select gene fusions, as previously described.(23, 24) To maximize mutation-detection accuracy (sensitivity and specificity) in impure clinical specimens, the test was previously optimized and validated to detect base substitutions at a ≥5% mutant allele frequency (MAF) and indels with a ≥10% MAF with ≥99% accuracy.(23, 24) The mutations reported on the F1 test were all manually, individually reviewed. Mutations were retained and included in the subsequent analysis only if either they were classified by Foundation Medicine as definitely 'known', were frameshift, nonsense or splice-site mutations or if a mutation affecting the same amino acid residue had been previously recognised in another tumour (determined by analysis of Catalogue of Somatic Mutations in Cancer (COSMIC) datasets directly (20) and via visualisation using the St Jude's PeCan data portal. (25))

Due to the greater depth of the FoundationOne Heme® test a cut off for mutations being present in ≥5% reads was applied to this dataset (mutations that could have been detected using the depth achieved in the MyXI study). Epigenetic genes from the list in **Supplementary Table 2** (and analysed in the MyXI data) which were also sequenced in the UAMS dataset are shown in **Supplementary Table 4.** This comprises: Core/linker histones (n=12), histone demethylases (n=5), histone methyltransferases (n=11), histone deacetylases (n=3), histone acetyltransferases (n=4), DNA modifiers (n=4), readers (n=2) and chromatin assembly/remodelling (n=9).

The percentage of patients with mutations in each class of epigenetic modifier were compared between the MyXI dataset and the UAMS dataset using the z-test, multiple testing was corrected for using the Bonferroni method. Since CCF could not be calculated for the patients in the UAMS dataset (due to the absence of copy

247 number data) the variant allele frequencies were compared for those genes with at
248 least 2 mutated samples in both datasets.

Results

Whole exome sequencing was performed on samples from 463 patients in the Myeloma XI trial prior to treatment. We identified mutations in genes encoding epigenetic modifiers in over half (53%) of patients. 20 epigenetic modifier genes were mutated in at least 5/463 (>1%) of individuals with frequencies shown in **Table 1,** cancer clonal fraction (CCF), shown in **Figure 1** and mutation location annotated in **Figure 2 and Supplementary Figure 1.** The distribution of the main translocation subgroups, known to drive myeloma pathogenesis, did not differ significantly between patients with and without a mutation in any epigenetic modifier (**Supplementary Figure 2**).

Below we report a detailed analysis of the frequency and potential clinical impact of mutations in each class of epigenetic modifier.

Core and linker histone mutations: HIST1H1E is significantly mutated at myeloma diagnosis

The gene *HIST1H1E,* which encodes a linker histone H1 protein, was found to be significantly mutated in the MyXI cohort(14) at diagnosis with mutations in 2.8% of patients (13/463, $p<1x10^{-10}$, $q<1x10^{-10}$). Further analysis revealed recurrent non-synonymous mutations in other histone 1 family genes including *HIST1H1B* (0.2%, 1/463), *HIST1H1C* (2.6%, 12/463), and *HIST1H1D* (0.6%, 3/463) with mutations in at least one of these genes (*HIST1H1B-E*) occurring in 6% (28/463) of patients. The mutations **(Figure 2a)** in these genes were predominantly missense SNVs affecting the globular domain of histone H1. HIST1H1B-E were aligned **(Figure 2b)** with sites affected by mutation highlighted. There were 3 sites of recurrent mutations at residues equivalent to alanine 61 and 65 and lysine 81 of HIST1H1E. There were

also several additional mutations within the globular domain across variants located between residues 100-110. The globular domain of these genes is a region frequently mutated in other cancers including follicular lymphoma and diffuse large B cell lymphoma.(26, 27) . The observations that the mutations occurred in a conserved region and there were no mutations in one known gene, *HIST1H1A,* supports the hypothesis that these are not passenger mutations and may carry some significance to myeloma pathogenesis.

Mutations in *HIST1H1B-E* were tested for their impact on protein function using SIFT analysis. This analysis was possible in 29/31 mutations and 69% (20/29) were found to be damaging (**Supplementary Table 5**). *HIST1H1B-E* mutations did not have a prognostic impact (**Supplementary Figure 3**).

CCF analysis showed *HIST1H1E* and *HIST1H1C* mutations to be highly clonal (**Figure 1**) suggesting they are either acquired early in clonal development or selected for at progression to symptomatic disease, but nevertheless play an important role in myeloma pathogenesis in these patients.

Mutations were also seen in the core histone proteins 2A (4.5%, 21/463 patients, one frameshift mutation, remainder missense SNVs), 2B (3.7%, 17/463, one frameshift mutation, remainder missense SNVs), 3 (2.8%, 13/463, one frameshift, one nonsense, remainder missense SNVs) and 4 (1.7%, 8/463, all missense SNVs) but with a low frequency of mutation in each family member, with no individual gene being mutated in more than 4 patients. The total number of patients harbouring a mutation in any histone protein (linker or core) was 18% (83/463) with the presence or absence of any mutation having no effect on progression or survival.

Histone modifier mutations: Mutations/deletions in KDM6A/UTX may shorten overall survival from diagnosis.

a) Methylation modifiers

Potentially deleterious mutations in histone methyltransferase/demethylase enzymes were found in 24% (112/463) of patients, though the percentage with each gene mutated was low (**Figure 3**). The most frequently mutated gene family in the methyltransferases was *MLL/2/3/4/5* (7% of patients, 31/463. By gene: *MLL* 1.7%, 8/463, *MLL2* 1.3%, 6/463, *MLL3* 1.5%, 7/463 *MLL4* 1.5%, 7/463 and *MLL5* 1.1%, 5/463) but no recurrent mutations were seen across variants. Of 36 mutations (across the 31 patients) 29 were missense SNVs, 4 nonsense SNVs, 2 splice site SNVs and 1 frameshift mutation. As in other diseases mutations in MLL family genes are widely distributed across the genes with no conserved sites or regions of mutation (**Supplementary Figure 1**). The majority of mutations identified in our patients were novel but those previously identified in other diseases included one in *MLL3*, p.R190Q (endometrial) (20) and in *MLL4*, p.R297* (large intestine) (28). The presence of MLL family mutations in myeloma patients did not have an effect on progression-free or overall survival.

No mutations in the H3K27 methyltransferase *EZH2* were detected, in contrast to the finding of recurrent mutations in other B cell malignancies. There were only 2 patients with mutations affecting the H3K36 methyltransferase MMSET, with none of the *MMSET* activating mutations at p.E1099K (previously identified in MM1.S myeloma cell line)(29) seen. *EHMT2*, the gene encoding the H3K9 methyltransferase G9a was mutated in 5/463 patients (1%, all missense mutations) with one in the SET domain (p.Y1097F) and one in the ankyrin repeat 'reader' domain (p.E699K) (**Supplementary Figure 1**).

The most frequently mutated demethylase gene was *KDM3B,* 1.5% (7/463) of patients with two splice site mutations, one nonsense mutation and 4 missense mutations, none of which have been previously identified in other cancers in the COSMIC database. The primary site of action of KDM3B is H3K9 and deletions of this gene have been implicated in myelodysplastic syndrome associated with 5q-. (30, 31)

KDM6A/UTX mutations were seen in 1.3% (6/463) of patients, were all missense mutations and were highly clonal (**Figure 1**). p.R118K, had been previously identified in lung cancer (32), p.Q398H in gastric adenocarcinoma (20) whilst another p.G66A occurred at the same residue as the p.G66D mutation previously identified in the OPM-2 myeloma cell line (10). Further analysis of this gene for whole exonic deletions increased the number of patients affected by a potentially inactivating lesion to 3% (15/463). Patients with a *KDM6A* mutation or deletion had a reduced overall survival (OS) compared to wild type on univariate analysis (**Figure 4A**) (medians not reached, logrank p=0.0498, percent alive at 2 years 51% 95%CI (30, 85) vs 80% 95%CI (77, 84) with a similar trend for progression-free survival (PFS).

b) Acetylation modifiers

The most frequently mutated gene encoding a histone acetyltransferase (HAT) or deacetylase (HDAC) was the HAT *EP300* mutated in 1.3% (6/463) of patients. Mutations in this gene have been previously identified in a number of other malignancies including DLBCL, where they are most commonly found in the HAT enzymatic domain.(33-35) One frameshift deletion (p.S90fs) and one nonsense mutation (p.Q1077*) are likely to result in an absence of functional protein. Of the 4 missense mutations one occurs in the active HAT domain (p.P1388S) and so may

directly affect the catalytic function of the protein, whilst one in the bromodomain (p.V1079L) and two in the zinc finger binding domain (p.T1775P, p.G1778P) may still have significant effects on the protein function by affecting target binding. Mutations in CREBBP, a closely related member of the KAT3 family of histone acetyltransferases were also found in 3 patients. One mutation, p.R1360Q occurs at the same amino acid as the p.R1360* mutation seen in several diseases including B-NHL (33). EP300/CREBBP have a wide range of targets and are able to acetylate all four histones as well as being involved in many cellular processes linked to cancer such as cell cycle progression, p53 activity, DNA repair and apoptosis.(36)

There were no genes encoding histone reader proteins that were mutated in more than 1% (5/463) of patients at diagnosis.

DNA methylation modifier mutations: Mutations in DNA methylation modifiers are associated with a shorter overall survival from diagnosis.

DNA methylation modifiers were mutated in 4% (17/463) of patients at diagnosis. This included specific mutations previously reported in numerous tumour types such as p.R132C in *IDH1* previously reported in glioma, chondrosarcoma and AML (37-41), and p.R140W in *IDH2* and p.C1378Y and p.Y1661* in *TET2* previously reported in AML/MDS (42-44) and p.E784K in DNMT3A previously reported in biliary tract tumours (45). A mutation was also seen at p.E477K in DNMT3A, the site of recurrent mutation in AML and MPNs (p.E477* and p.E477fs)(46-48). Collectively, mutations in any DNA methylation modifier (*TET1/2/3* n=11, *IDH1/2* n=2 or *DNMT1/3A/B* n=6) were associated with a shorter OS on univariate analysis, **Figure 4B** (medians not reached, p=0.045, % alive at 2 years 58% 95%CI (39, 88) vs 80% 95%CI (76, 84). There was no significant effect of PFS. This effect on OS with no effect on PFS is explained by a significantly shortened post-progression survival for

patients with mutations in DNA methylation modifiers compared to those without (p=0.002, logrank test, data not shown).

Chromatin remodelling complex mutations

The most frequently mutated genes involved in chromatin remodelling were *CHD4* in 1.9% (9/463, all missense) of patients and *CHD2* in 1.5% (7/463, one frameshift, six missense). Both are members of the nucleosome remodelling and deacetylase complex (NuRD). This highlights the potential role of the NuRD complex in myeloma pathogenesis however its action may not be entirely epigenetic as it has also been recently noted that CHD4 may also have roles in DNA damage repair and cell cycle progression independent of the NuRD complex and epigenetic activity.(49)

Other remodelling genes mutated in more than 1% of patients include *ARID1A* (encoding BAF 250a) in 1.3% (6/463, 2 nonsense and 4 missense) and *ARID2* (encoding BAF200) in 1.3% (6/463, 3 nonsense, 1 splice site, 2 missense). Their gene products are part of in the SWI/SNF (sucrose non-fermenting/ switch non-fermenting) remodelling complex and are responsible for its interaction with DNA. Mutations in genes encoding members of this complex have been previously demonstrated to be recurrently mutated in both solid tumours and haematologic malignancies in up to 20% of cancer patients overall.(50, 51) The mechanisms behind this have yet to be fully elucidated and in different diseases has been suggested to be related to the role of the SWI/SNF complex in DNA damage repair,(52) nucleosome positioning, DNAse hypersensitivity sites, the regulation of developmental gene expression and/or the interaction of the complex with both histone and DNA modifiers.(53) We found previously identified mutations at p.M918I in ARID1A (seen in renal cell carcinoma) (20) at p.Q937* in ARID2 (seen in biliary

tract carcinoma and melanoma) (20, 45), p.Q1611* in ARID2 (seen in squamous cell carcinoma) (54) and p.A1555S in ARID2 (seen recurrently in head and neck carcinoma) (55).

Survival analysis

For each class of epigenetic modifier the progression-free and overall survival for those patients with a mutation in any gene within the class were compared to those patients without. In addition those patients with mutations of interest were compared to those without in the following cases, *KDM6A* mutations or deletions, MLL family gene mutations and Histone 1 family gene mutations. An effect on overall survival was found on univariate analysis for *KDM6A* mutations/deletions and DNA modifier mutations as described above. Multivariate analysis using a cox-regression model was therefore carried out considering other factors known to have an adverse effect on overall survival; presence of an adverse translocation t(4;14), t(14;16) or t(14;20), del17p, gain or amplification of 1q, international staging score and age >70. In this model DNA modifier mutations, in addition to del17p, ISS and gain or amplification of 1q remained significant but not KDM6A mutations/deletions. (**Supplementary Table 6**).

The frequency of mutations in epigenetic modifiers increases following treatment

Longitudinal investigations are planned for the patients in the UK Myeloma XI trial, but at present the majority remain in remission. Therefore, in order to study the likely differences in the frequency of mutations in epigenetic modifiers as disease progresses we utilised available data for a series of 156 previously treated myeloma patients from the University of Arkansas for Medical Sciences (UAMS) who underwent targeted sequencing. Due to the different methods used we restricted our

comparison to only those epigenetic modifier genes that have been sequenced in both studies (**Supplementary Table 4**) Results are summarised in **Figure 5A** and **Supplementary Table 7** and show an increase in the number of patients with a mutation in any epigenetic modifier in samples taken at later stages of disease. There is a statistically significant increase in the number of patients with a mutation in any histone methyltransferase gene (6.9%, 32/463 MyXI vs 17%, 26/156 UAMS) or any DNA methylation modifier (1.9%, 9/463 MyXI vs 8.3%, 13/156 UAMS) and a notable increase in histone acetyl-transferase gene mutations (2.4%, 11/463 MyXI vs 7.1%, 11/156 UAMS). These changes appear to be the result of increases in mutations in *MLL2*, *MLL3*, *SETD2*, *CREBBP*, *DNMT3A* and *TET2* (**Figure 5B**).

The variant allele frequency was compared between the MYXI and UAMS samples (**Supplementary Figure 4)** with no statistically significant differences seen. An analysis of the distribution of mutations across risk groups, as defined by gene expression profile risk score (GEP70) **(Supplementary Figure 5)** and UAMS molecular subgroups (**Supplementary Figure 6)** found a slight overrepresentation of the PR subgroup in those patients with an epigenetic modifier mutation compared to those without.

Of note in the UAMS dataset 2 of the 3 mutations in *HIST1H1E* had been previously identified, p.A65P in the MyXI dataset and p.P161S in a lymphoid neoplasm (COSMIC) whilst the third, p.A47V, occurred at the same residue as p.A47P seen in a MyXI patient. This supports the evidence of likely impact of *HIST1H1E* mutations in myeloma suggested by the MyXI analysis. Recurrent mutations were also seen in the UAMS dataset at the same location in IDH1, p.R132C and p.R132H, with p.R132C having been also identified in a MyXI patient.

Discussion

We report mutations within genes encoding epigenetic modifiers in myeloma with an impact on pathology and survival at diagnosis and an increased frequency after treatment. The spectrum of mutations in myeloma is broad with no single epigenetic modifier being mutated in a large proportion of patients. Several of the mutations identified have been previously related to cancer pathogenesis and/or may open possibilities for targeted treatment strategies for subgroups of patients. This work changes our understanding of the epigenetic landscape of myeloma exposing a wider spectrum of epigenetic processes than previously recognised, that may be altered in large numbers of patients, affecting disease biology and outcome.

One of our key findings is the significant mutation of *HIST1H1E* and similar mutations across Histone 1 family genes in diagnostic samples. Histone H1 acts to control the higher order structure of chromatin by spacing nucleosomes and holding DNA in place as it winds around each nucleosome octamer. It has, therefore, been suggested to affect gene transcription via the modulation of the accessibility of DNA to transcription factors. Mutations in the histone H1 family genes have been found in other haematological malignancies including follicular lymphoma and diffuse large B cell lymphoma.(27, 56) In follicular lymphoma these were also noted to be predominantly in the globular domain and one such mutation (Ser102Phe) was functionally demonstrated to affect impaired ability of histone H1 to associate with chromatin.(56) This is close to several sites of mutation identified in our study between residues 100-110. More recently, histone H1 has also been shown to play a role in regulating DNA methylation via DNMT1 and DNMT3B binding and altering H3K4 methylation by affecting binding of the methyltransferases SET7/9.(26, 57) It

might, therefore, be postulated that it is via these mechanisms that mutations in histone H1 may have an oncogenic effect. Analysis of paired sample mutation and DNA methylation analysis will further inform this hypothesis.

HIST1H1E and *HIST1H1C* mutations (along with *KDM6A* and *ARID2*) were highly clonal, suggesting they may be acquired early in pathogenesis or selected for at progression to symptomatic disease. Our analysis estimates CCF using a method that correlates well with single cell analysis results; however in certain situations copy number alterations or polyclonality may yield anomalous CCF estimates for individual mutations. Future improvements in techniques for calculating CCF may further inform these results.

Our analysis identifies 3% of patients at diagnosis with a potentially inactivating lesion in *KDM6A/UTX*. This is lower than the frequency of mutations previously identified (10) and often reported of 6/58, 10%. This previous analysis, however, included cell lines, and when these are removed and only patient samples from the analysis considered the percentage with a lesion falls to 4% (2/49),in keeping with our study.

The Myeloma XI study pairs mutation and outcome data giving us the first opportunity to explore any possible association between epigenetic modifier mutations and outcome. Individual epigenetic modifier genes are mutated in small numbers of patients and so the size of our dataset limits the power to detect any specific gene effects. We therefore grouped mutations (11 groups as defined above) and on univariate analysis identified a detrimental effect on overall survival of KDM6A mutation/deletion (Log-rank p=0.0498) and DNA methylation modifiers (Log-rank p=0.045). If a Bonferroni correction for multiple testing were applied to this data

the p-values obtained would fall above the level considered significant, however this arbitrary cut off may miss a clinically meaningful effect. Further investigation in future studies will help to clarify this. Notably the effect of DNA methylation modifier mutations on overall survival withstood multivariate analysis and in larger cohorts it would be of interest to investigate the independent association of mutations within this, and other groups, on outcomes.Myeloma is part of a spectrum of malignancies arising from B cell populations at various stages of B cell ontogeny. Mutations in epigenetic modifiers are seen across this spectrum but we can now show that different patterns are seen dependent upon the biology of the population examined. Recurrent mutations in *HIST1H1*, *MLL* and *EZH* gene families have been identified in diffuse large B cell lymphoma and follicular lymphoma.(26, 34) We show that *HIST1H1* and *MLL* mutations are seen in myeloma, although at a much lower frequency, whereas *EZH2* mutations are not seen, suggesting a different pathogenic mechanism. The different spectrum of epigenetic mutations is more marked when we compare lymphoid to myeloid diseases such as acute myeloid leukaemia, myelodysplastic syndromes and myeloproliferative neoplasms, where mutations in DNA modifying enzymes such as *DNMT3A*, *IDH* and *TET2* predominate.(58, 59) We found mutations in these genes in myeloma, but at a much lower frequency.

Our results highlight possible targeted treatment approaches for patients either at diagnosis or at relapse. For example patients with a *KDM6A/UTX* mutation or deletion might be amenable to the use of EZH2 inhibitors, currently in the early stages of clinical development for lymphoma patients. Inhibiting EZH2, the H3K27 methyltransferase, may counteract the increased H3K27 methylation resulting from inactivation of the demethylase. A recent study also reports that *ARID1A* mutated cancers may be sensitive to EZH2 inhibition, demonstrating a synthetic lethal effect

via the PI3-AKT pathway.(60) Patients with *IDH* mutations might be amenable to IDH inhibitors currently in early stages of development and a more global strategy might be possible for patients with mutations in DNA methylation modifiers with demethylating agents.

On comparison with results from focused sequencing of 156 previously treated patients we show an increase in the number of patients with a mutation in genes encoding a histone methyltransferases and DNA methylation modifiers. This suggests that these events may either play a role in disease progression or occur more frequently following exposure to induction chemotherapy in resistant subclones. The change in frequency of mutations in DNA methylation modifiers as myeloma progresses is supported by data showing a change in the methylation pattern at different disease stages.(3) There are several limitations to our comparison, however, including the different sequencing methods and depth. These results will, therefore, need to be validated in matched patient populations following relapse within a clinical trial setting using the same analysis method.

This whole-exome analysis of Myeloma XI patients at diagnosis is the first extensive analysis of the spectrum of mutations in epigenetic modifiers in a uniformly treated population in myeloma. An association with clinical outcome for *KDM6A* mutated or deleted patients, and mutations in DNA methylation modifiers is suggested in our dataset but will need validation in larger studies or meta-analysis due to the low overall frequency of the mutations. These data further emphasise the importance of epigenetics in myeloma and provide potential new targets for personalised therapeutic strategies for patients. Our findings support the use of sequencing-based diagnostics in myeloma both at diagnosis and relapse in order to identify potentially prognostic and/or targetable lesions.

549

Acknowledgements:

We would like to thank all the patients and staff at centres throughout the UK whose participation made this study possible. We are grateful to the NCRI Haemato-oncology subgroup and to all principle investigators for their dedication and commitment to recruiting patients to the study. The principal investigators at the four top recruiting centres were Dr Don Milligan (Heart of England NHS Foundation Trust), Dr Jindriska Lindsay (Kent and Canterbury Hospital), Dr Nigel Russell (Nottingham University Hospital) and Dr Clare Chapman (Leicester Royal Infirmary). The support of the Clinical Trials Research Unit at The University of Leeds was essential to the successful running of the study and the authors would like to thank all the staff including Helen Howard, Corrine Collett, Jacqueline Ouzman and Alex Szubert. We also acknowledge The Institute of Cancer Research Tumour Profiling Unit for their support and technical expertise in this study. Thanks to Tim Vojt for help with figure 3.

Grant support:

This work was supported by a Myeloma UK program grant, Cancer Research UK CTAAC sample collection grants (C2470/A12136 and C2470/A17761), a Cancer Research UK Biomarkers and Imaging Discovery and Development grant (C2470/A14261), and funds from the National Institute of Health Biomedical Research Centre at the Royal Marsden Hospital and Institute of Cancer Research. Work carried out at the Myeloma Institute, University of Arkansas for Medical Sciences was supported by a grant from the NIH. CP is a Wellcome Trust Clinical Research Fellow. FED is a CRUK Senior Cancer Fellow.

Genome data from the Myeloma XI trial patients have been deposited at the European Genome-phenome Archive (EGA, http://www.ebi.ac.uk/ega/) which is hosted at the EBI, under accession number EGAS00001001147.

References

1. Morgan GJ, Walker BA, Davies FE. The genetic architecture of multiple myeloma. Nature reviews Cancer. 2012;12:335-48.
2. Pawlyn C, Kaiser MF, Davies FE, Morgan GJ. Current and potential epigenetic targets in multiple myeloma. Epigenomics. 2014;6:215-28.
3. Walker BA, Wardell CP, Chiecchio L, Smith EM, Boyd KD, Neri A, et al. Aberrant global methylation patterns affect the molecular pathogenesis and prognosis of multiple myeloma. Blood. 2011;117:553-62.
4. Ezponda T, Popovic R, Shah MY, Martinez-Garcia E, Zheng Y, Min DJ, et al. The histone methyltransferase MMSET/WHSC1 activates TWIST1 to promote an epithelial-mesenchymal transition and invasive properties of prostate cancer. Oncogene. 2013;32:2882-90.
5. Hudlebusch HR, Theilgaard-Monch K, Lodahl M, Johnsen HE, Rasmussen T. Identification of ID-1 as a potential target gene of MMSET in multiple myeloma. British journal of haematology. 2005;130:700-8.
6. Kuo AJ, Cheung P, Chen K, Zee BM, Kioi M, Lauring J, et al. NSD2 links dimethylation of histone H3 at lysine 36 to oncogenic programming. Molecular cell. 2011;44:609-20.
7. Martinez-Garcia E, Popovic R, Min DJ, Sweet SM, Thomas PM, Zamdborg L, et al. The MMSET histone methyl transferase switches global histone methylation and alters gene expression in t(4;14) multiple myeloma cells. Blood. 2011;117:211-20.
8. Keats JJ, Reiman T, Maxwell CA, Taylor BJ, Larratt LM, Mant MJ, et al. In multiple myeloma, t(4;14)(p16;q32) is an adverse prognostic factor irrespective of FGFR3 expression. Blood. 2003;101:1520-9.
9. Walker BA, Leone PE, Chiecchio L, Dickens NJ, Jenner MW, Boyd KD, et al. A compendium of myeloma-associated chromosomal copy number abnormalities and their prognostic value. Blood. 2010;116:e56-65.
10. van Haaften G, Dalgliesh GL, Davies H, Chen L, Bignell G, Greenman C, et al. Somatic mutations of the histone H3K27 demethylase gene UTX in human cancer. Nature genetics. 2009;41:521-3.
11. Chapman MA, Lawrence MS, Keats JJ, Cibulskis K, Sougnez C, Schinzel AC, et al. Initial genome sequencing and analysis of multiple myeloma. Nature. 2011;471:467-72.
12. Oyer JA, Huang X, Zheng Y, Shim J, Ezponda T, Carpenter Z, et al. Point mutation E1099K in MMSET/NSD2 enhances its methyltranferase activity and leads to altered global chromatin methylation in lymphoid malignancies. Leukemia. 2014;28:198-201
13. Walker BA, Wardell CP, Murison A, Boyle EM, Begum DB, Dahir NM, et al. APOBEC family mutational signatures are associated with poor prognosis translocations in multiple myeloma. Nature communications. 2015;6:6997.
14. Walker BA, Boyle EM, Wardell CP, Murison A, Begum DB, Dahir NM, et al. Mutational Spectrum, Copy Number Changes, and Outcome: Results of a Sequencing Study of Patients With Newly Diagnosed Myeloma. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2015;33:3911-20.
15. Kozarewa I, Rosa-Rosa JM, Wardell CP, Walker BA, Fenwick K, Assiotis I, et al. A modified method for whole exome resequencing from minimal amounts of starting DNA. PloS one. 2012;7:e32617.
16. Boeva V, Popova T, Bleakley K, Chiche P, Cappo J, Schleiermacher G, et al. Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. Bioinformatics. 2012;28:423-5.
17. Boyle EM, Proszek PZ, Kaiser MF, Begum D, Dahir N, Savola S, et al. A molecular diagnostic approach able to detect the recurrent genetic prognostic factors typical of presenting myeloma. Genes, chromosomes & cancer. 2015;54:91-8
18. Melchor L, Brioli A, Wardell CP, Murison A, Potter NE, Kaiser MF, et al. Single-cell genetic analysis reveals the composition of initiating clones and phylogenetic patterns of branching and parallel evolution in myeloma. Leukemia. 2014;28:1705-15

19. St. Jude Children's Research Hospital - Washington University. Pediatric Cancer Genome Project.
20. Catalogue of Somatic Mutations in Cancer (COSMIC) database accessed at http://www.cancer.sanger.ac.uk Dec 2015 (COSMIC data IDs 419, 540, 541, 588 have no other references).
21. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nature protocols. 2009;4:1073-81.
22. Heuck CJ, Qu P, van Rhee F, Waheed S, Usmani SZ, Epstein J, et al. Five gene probes carry most of the discriminatory power of the 70-gene risk model in multiple myeloma. Leukemia. 2014;28:2410-3.
23. Frampton GM, Fichtenholtz A, Otto GA, Wang K, Downing SR, He J, et al. Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. Nature biotechnology. 2013;31:1023-31.
24. Lipson D. Identification Of Actionable Genomic Alterations In Hematologic Malignancies By a Clinical Next Generation Sequencing-Based Assay (Abstract). Blood. 2013;122:230-.
25. Zhou X, Edmonson MN, Wilkinson MR, Patel A, Wu G, Liu Y, et al. Exploring genomic alteration in pediatric cancer using ProteinPaint. Nat Genet. 2016;48:4-6.
26. Li H, Kaminski MS, Li Y, Yildiz M, Ouillette P, Jones S, et al. Mutations in linker histone genes HIST1H1 B, C, D, and E; OCT2 (POU2F2); IRF8; and ARID1A underlying the pathogenesis of follicular lymphoma. Blood. 2014;123:1487-98.
27. Morin RD, Mungall K, Pleasance E, Mungall AJ, Goya R, Huff RD, et al. Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. Blood. 2013;122:1256-65.
28. Giannakis M, Hodis E, Jasmine Mu X, Yamauchi M, Rosenbluh J, Cibulskis K, et al. RNF43 is frequently mutated in colorectal and endometrial cancers. Nature genetics. 2014;46:1264-6.
29. Oyer JA, Huang X, Zheng Y, Shim J, Ezponda T, Carpenter Z, et al. Point mutation E1099K in MMSET/NSD2 enhances its methyltranferase activity and leads to altered global chromatin methylation in lymphoid malignancies. Leukemia. 2014;28:198-201.
30. Graubert TA, Payton MA, Shao J, Walgren RA, Monahan RS, Frater JL, et al. Integrated genomic analysis implicates haploinsufficiency of multiple chromosome 5q31.2 genes in de novo myelodysplastic syndromes pathogenesis. PloS one. 2009;4:e4583.
31. Hu Z, Gomes I, Horrigan SK, Kravarusic J, Mar B, Arbieva Z, et al. A novel nuclear protein, 5qNCA (LOC51780) is a candidate for the myeloid leukemia tumor suppressor gene on chromosome 5 band q31. Oncogene. 2001;20:6946-54.
32. Imielinski M, Berger AH, Hammerman PS, Hernandez B, Pugh TJ, Hodis E, et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. Cell. 2012;150:1107-20.
33. Pasqualucci L, Dominguez-Sola D, Chiarenza A, Fabbri G, Grunn A, Trifonov V, et al. Inactivating mutations of acetyltransferase genes in B-cell lymphoma. Nature. 2011;471:189-95.
34. Pasqualucci L, Trifonov V, Fabbri G, Ma J, Rossi D, Chiarenza A, et al. Analysis of the coding genome of diffuse large B-cell lymphoma. Nature genetics. 2011;43:830-7.
35. Morin RD, Mendez-Lago M, Mungall AJ, Goya R, Mungall KL, Corbett RD, et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. Nature. 2011;476:298-303.
36. Iyer NG, Ozdag H, Caldas C. p300/CBP and cancer. Oncogene. 2004;23:4225-31.
37. Bleeker FE, Lamba S, Leenstra S, Troost D, Hulsebos T, Vandertop WP, et al. IDH1 mutations at residue p.R132 (IDH1(R132)) occur frequently in high-grade gliomas but not in other solid tumors. Human mutation. 2009;30:7-11.
38. Schnittger S, Haferlach C, Ulke M, Alpermann T, Kern W, Haferlach T. IDH1 mutations are detected in 6.6% of 1414 AML patients and are associated with intermediate risk karyotype and unfavorable prognosis in adults younger than 60 years and unmutated NPM1 status. Blood. 2010;116:5486-96.

686 39. Mardis ER, Ding L, Dooling DJ, Larson DE, McLellan MD, Chen K, et al. Recurring
687 mutations found by sequencing an acute myeloid leukemia genome. The New England
688 journal of medicine. 2009;361:1058-66.
689 40. Marcucci G, Maharry K, Wu YZ, Radmacher MD, Mrozek K, Margeson D, et al. IDH1
690 and IDH2 gene mutations identify novel molecular subsets within de novo cytogenetically
691 normal acute myeloid leukemia: a Cancer and Leukemia Group B study. Journal of clinical
692 oncology : official journal of the American Society of Clinical Oncology. 2010;28:2348-55.
693 41. Amary MF, Bacsi K, Maggiani F, Damato S, Halai D, Berisha F, et al. IDH1 and IDH2
694 mutations are frequent events in central chondrosarcoma and central and periosteal
695 chondromas but not in other mesenchymal tumours. The Journal of pathology.
696 2011;224:334-43.
697 42. Jeziskova I, Razga F, Bajerova M, Racil Z, Mayer J, Dvorakova D. IDH2 mutations in
698 patients with acute myeloid leukemia: missense p.R140 mutations are linked to disease
699 status. Leukemia & lymphoma. 2010;51:2285-7.
700 43. Andersson AK, Miller DW, Lynch JA, Lemoff AS, Cai Z, Pounds SB, et al. IDH1 and
701 IDH2 mutations in pediatric acute leukemia. Leukemia. 2011;25:1570-7.
702 44. Papaemmanuil E, Gerstung M, Malcovati L, Tauro S, Gundem G, Van Loo P, et al.
703 Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood.
704 2013;122:3616-27; quiz 99.
705 45. Li M, Zhang Z, Li X, Ye J, Wu X, Tan Z, et al. Whole-exome and targeted gene
706 sequencing of gallbladder carcinoma identifies recurrent mutations in the ErbB pathway.
707 Nature genetics. 2014;46:872-6.
708 46. Gaidzik VI, Schlenk RF, Paschka P, Stolzle A, Spath D, Kuendgen A, et al. Clinical
709 impact of DNMT3A mutations in younger adult patients with acute myeloid leukemia: results
710 of the AML Study Group (AMLSG). Blood. 2013;121:4769-77.
711 47. Ley TJ, Ding L, Walter MJ, McLellan MD, Lamprecht T, Larson DE, et al. DNMT3A
712 mutations in acute myeloid leukemia. The New England journal of medicine. 2010;363:2424-
713 33.
714 48. Stegelmann F, Bullinger L, Schlenk RF, Paschka P, Griesshammer M, Blersch C, et
715 al. DNMT3A mutations in myeloproliferative neoplasms. Leukemia. 2011;25:1217-9.
716 49. O'Shaughnessy A, Hendrich B. CHD4 in the DNA-damage response and cell cycle
717 progression: not so NuRDy now. Biochemical Society transactions. 2013;41:777-82.
718 50. Kadoch C, Hargreaves DC, Hodges C, Elias L, Ho L, Ranish J, et al. Proteomic and
719 bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in
720 human malignancy. Nature genetics. 2013;45:592-601.
721 51. Shain AH, Pollack JR. The spectrum of SWI/SNF mutations, ubiquitous in human
722 cancers. PloS one. 2013;8:e55119.
723 52. Watanabe R, Ui A, Kanno S, Ogiwara H, Nagase T, Kohno T, et al. SWI/SNF factors
724 required for cellular resistance to DNA damage include ARID1A and ARID1B and show
725 interdependent protein stability. Cancer research. 2014;74:2465-75.
726 53. Skulte KA, Phan L, Clark SJ, Taberlay PC. Chromatin remodeler mutations in human
727 cancers: epigenetic implications. Epigenomics. 2014;6:397-414.
728 54. Pickering CR, Zhou JH, Lee JJ, Drummond JA, Peng SA, Saade RE, et al.
729 Mutational landscape of aggressive cutaneous squamous cell carcinoma. Clinical cancer
730 research : an official journal of the American Association for Cancer Research.
731 2014;20:6582-92.
732 55. Martin D, Abba MC, Molinolo AA, Vitale-Cross L, Wang Z, Zaida M, et al. The head
733 and neck cancer cell oncogenome: a platform for the development of precision molecular
734 therapies. Oncotarget. 2014;5:8906-23.
735 56. Okosun J, Bodor C, Wang J, Araf S, Yang CY, Pan C, et al. Integrated genomic
736 analysis identifies recurrent mutations and evolution patterns driving the initiation and
737 progression of follicular lymphoma. Nature genetics. 2014;46:176-81.
738 57. Yang SM, Kim BJ, Norwood Toro L, Skoultchi AI. H1 linker histone promotes
739 epigenetic silencing by regulating both DNA methylation and histone H3 methylation.

Proceedings of the National Academy of Sciences of the United States of America. 2013;110:1708-13.
58. Haferlach T, Nagata Y, Grossmann V, Okuno Y, Bacher U, Nagae G, et al. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. Leukemia. 2014;28:241-7.
59. Im AP, Sehgal AR, Carroll MP, Smith BD, Tefferi A, Johnson DE, et al. DNMT3A and IDH mutations in acute myeloid leukemia and other myeloid malignancies: associations with prognosis and potential treatment strategies. Leukemia. 2014;28:1774-83.
60. Bitler BG, Aird KM, Garipov A, Li H, Amatangelo M, Kossenkov AV, et al. Synthetic lethality by targeting EZH2 methyltransferase activity in ARID1A-mutated cancers. Nature medicine. 2015;21:231-8.

Figure Legends

Figure 1

Cancer Clonal Fraction in genes encoding epigenetic modifiers.

The proportion of tumour cells containing an SNV was calculated in the 370 samples with accurate copy number assessments and plotted as the CCF interval for each sample with the indicated mutation, corrected for coverage. Mutations were considered clonal and shown in red if the upper CCF interval was ≥ 95% and sub-clonal if the upper CCF interval was <95%. The proportion of samples with each gene mutation that were clonal and sub-clonal are shown in the bar chart below the gene name. Epigenetic modifiers were analysed for CCF if they were mutated in ≥5/463 (>1%) patients.

Figure 2

Site of mutations in the genes encoding histone 1 variants

A, The common protein variants are shown with the globular domain of each highlighted in purple. The sites of mutation are indicated by the markers with red dots indicating missense mutations and purple dots frameshift mutations. There was 1 mutation in *HIST1H1B* in 1 patient, 12 mutations in *HIST1H1C* in 12 patients, 3 mutations in *HIST1H1D* in 3 patients and 15 mutations in *HIST1H1E* in 13 patients. There were no mutations in *HIST1H1A*.

B, The common proteins of Histone 1 are aligned and labelled such that 1H1A/H1.1 indicates the protein histone 1.1, encoded by gene *HIST1H1A,* 1H1B/H1.5 indicates the protein histone 1.5 encoded by gene *HIST1H1B* etc. The numbers indicate the amino acid number. The protein domains are indicated by the coloured bars overlying the protein sequence, green denotes the N terminal domain, purple the globular domain and red the C terminal domain. Mutations found in the patients sequenced are indicated by coloured square, pink = missense SNV, yellow = frame shift insertion and green = frame shift deletion. At some sites there was more than one patient with a mutation. The overlying arrows indicate amino acids where the mutations in different protein variants occur at the same equivalent amino acid residues.

Figure 3

Mutations in histone methyltransferases and demethylases.

The methyltransferases and demethylases located at their most commonly recognised site of activity on histone 3 are shown with the percentage of patients harbouring a mutation. Those in bold are mutated in 5 or more patients in the dataset.

Figure 4

Overall survival is shorter in patients with a *KDM6A* mutation or deletion and those with a mutation in any DNA modifier.

A, Kaplan-Meier curves showing patients with a *KDM6A* mutation or deletion (n=15, dashed) and those without (n=448, solid). Progression-free survival median mut/del 16.8 months vs wild type 26.6 months (logrank p=0.695). Overall survival medians not reached, (logrank p = 0.0498). Percent alive

at 2 years: mut/del 51% 95%CI (30, 85), wild type 80% 95%CI (77, 84). Survival on x-axis is plotted as the time since randomisation.

B, Kaplan-Meier curves showing patients with a DNA modifier mutation (n=17: *TET1/2/3* n=11, *IDH1/2* n=2 or *DNMT1/3A/B* n=6, dashed) and those without (n=446, solid). Progression-free survival median mut/del not reached vs wild type 26.6 months (logrank p=0.852). Overall survival medians not reached, (logrank p = 0.0455). Percent alive at 2 years: mut 58% 95%CI (39, 88), wild type 80% 95%CI (76, 84). Survival on x-axis is plotted as the time since randomisation.

Figure 5

The frequency of mutations in genes encoding epigenetic modifiers at presentation and following treatment by class of epigenetic modifier.

A, A comparison between the percentage of patients with a mutation in each class of epigenetic modifier at presentation (Myeloma XI data, shown in black bars) and after treatment (UAMS data, shown in gray bars). There is a significant increase in the percentage of previously treated patients with a mutation for those groups indicated by *. (z-test of difference in proportions $P<0.05$). Those comparisons remaining significant after multiple test correction (Boferroni method) are indicated by **. Full data is shown in supplementary table 7. Only mutations in genes sequenced in both studies (i.e. the genes sequenced in the UAMS F1 dataset, supplementary table 4) were included for this comparison.

B, The frequency (%) of mutations in genes of interest within each class

Table 1

Epigenetic genes mutated in ≥5/463 (>1%) newly diagnosed patients in Myeloma XI dataset

Gene	No of patients (total = 463)	Percentage
ARID1A	6	1.30%
ARID2	6	1.30%
ARID4A	5	1.08%
ARID5B	5	1.08%
CHD2	7	1.51%
CHD4	9	1.94%
EHMT2	5	1.08%
EP300	6	1.30%
HIST1H1C	12	2.59%
HIST1H1E	13	2.81%
KDM3B	7	1.51%
KDM6A	6	1.30%
MLL	8	1.73%
MLL2	6	1.30%
MLL3	7	1.51%
MLL4	7	1.51%
MLL5	5	1.08%
SETD2	6	1.30%
SETDB1	5	1.08%
TET2	5	1.08%

Figure 1

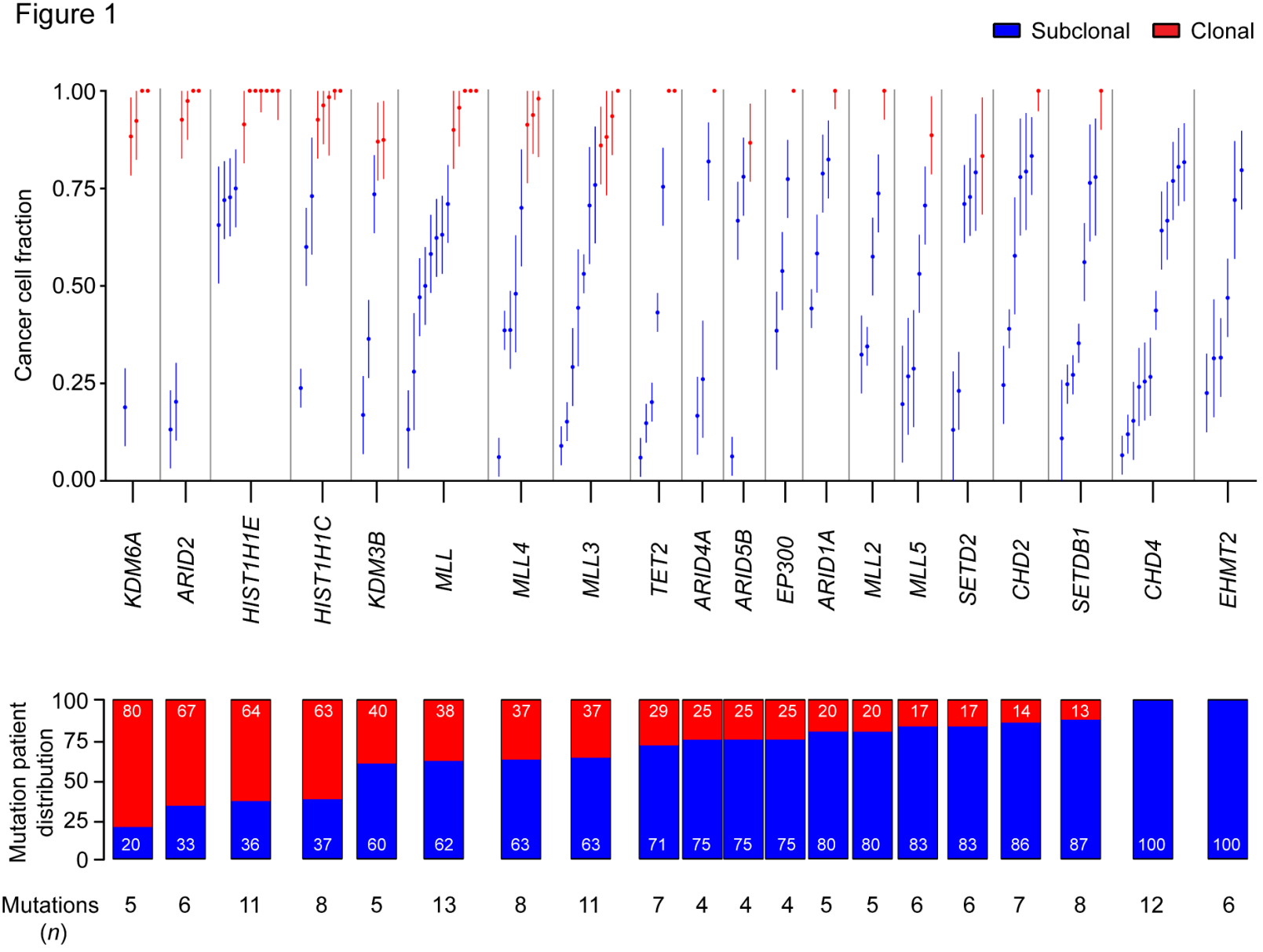
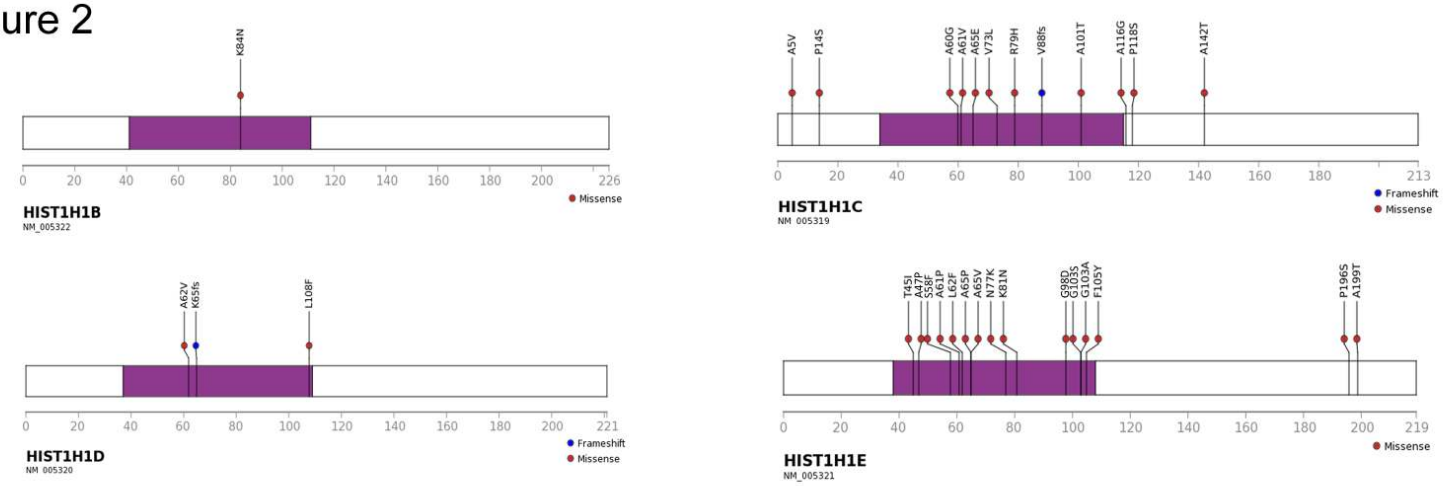

Subclonal
Clonal
Cancer cell fraction
1.00
0.75
0.50
0.25
0.00
KDM6A
ARID2
HIST1H1E
HIST1H1C
KDM3B
MLL
MLL4
MLL3
TET2
ARID4A
ARID5B
EP300
ARID1A
MLL2
MLL5
SETD2
CHD2
SETDB1
CHD4
EHMT2
Mutation patient distribution
100
75
50
25
0
80 67 64 63 40 38 37 37 29 25 25 25 20 20 17 17 14 13
20 33 36 37 60 62 63 63 71 75 75 75 80 80 83 83 86 87 100 100
Mutations (n)
5 6 11 8 5 13 8 11 7 4 4 4 5 5 6 6 7 8 12 6

Figure 2

a)


K84N
HIST1H1B
NM_005322
Missense
A5V
P14S
A60G
A61V
A65E
V73L
R79H
V88fs
A101T
A116G
P118S
A142T
HIST1H1C
NM_005319
Frameshift
Missense
A62V
K65fs
L108F
HIST1H1D
NM_005320
Frameshift
Missense
T45I
A47P
S58F
A61P
L62F
A65P
A65V
N77K
K81N
G98D
G103S
G103A
F105Y
P196S
A199T
HIST1H1E
NM_005321
Missense

b)

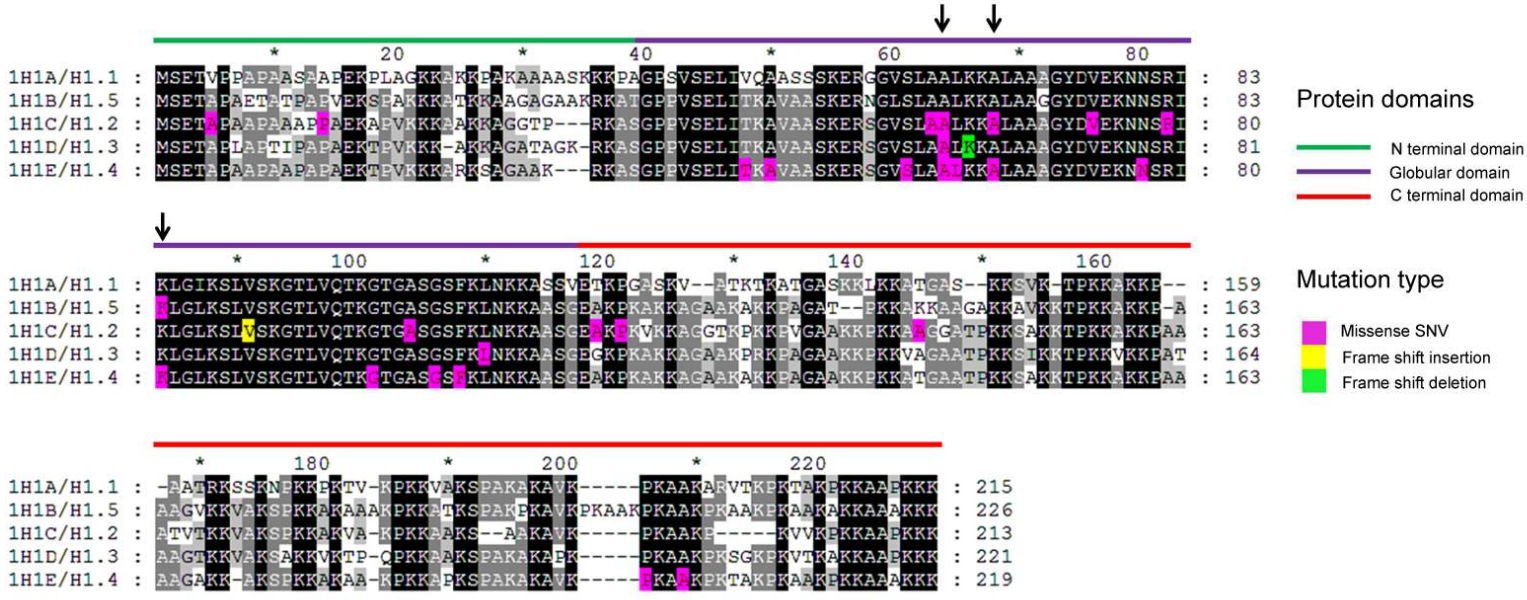

1H1A/H1.1
1H1B/H1.5
1H1C/H1.2
1H1D/H1.3
1H1E/H1.4
Protein domains
N terminal domain
Globular domain
C terminal domain
Mutation type
Missense SNV
Frame shift insertion
Frame shift deletion

Figure 3

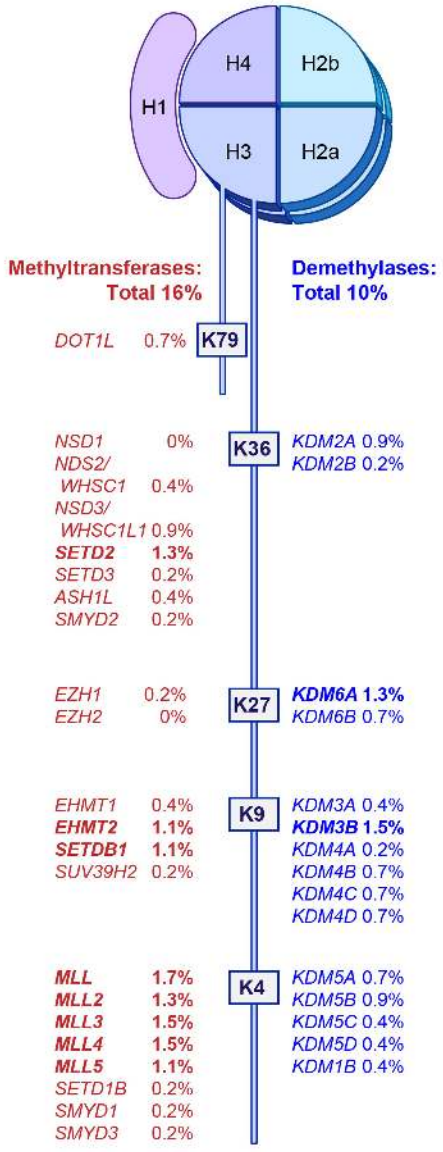
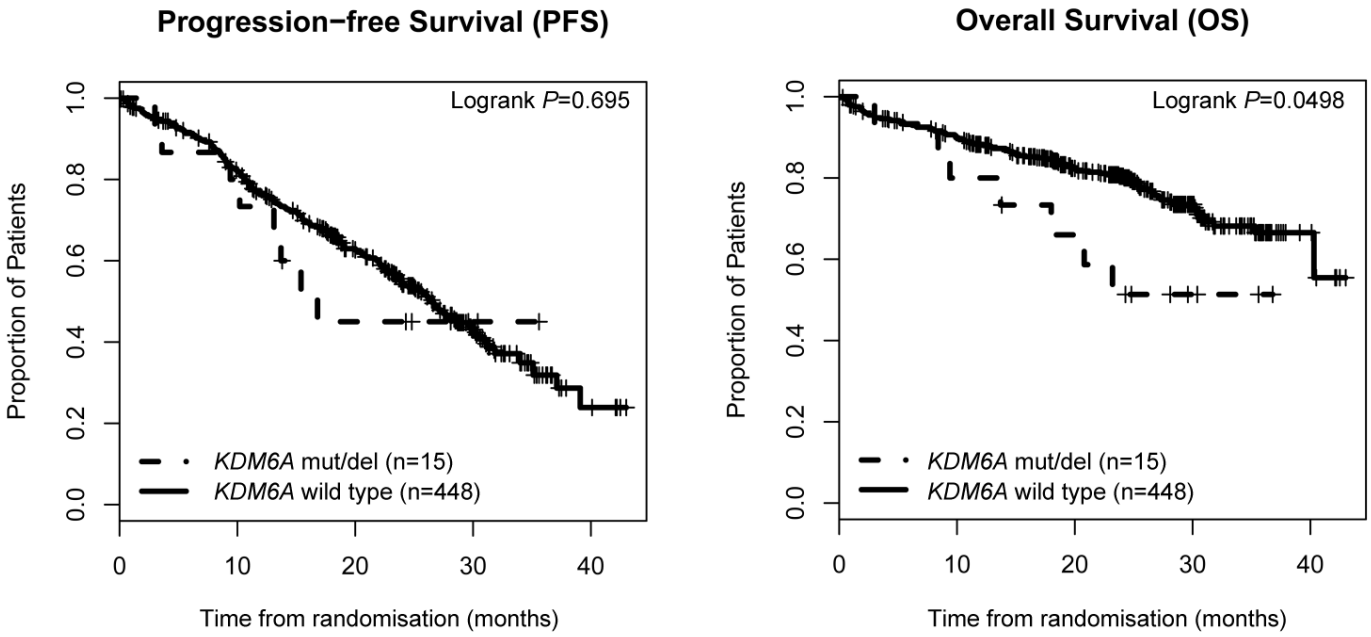

H4
H2b
H1
H3
H2a
Methyltransferases:
Total 16%
Demethylases:
Total 10%
DOT1L 0.7%
K79
NSD1 0%
NDS2/
WHSC1 0.4%
NSD3/
WHSC1L1 0.9%
SETD2 1.3%
SETD3 0.2%
ASH1L 0.4%
SMYD2 0.2%
K36
KDM2A 0.9%
KDM2B 0.2%
EZH1 0.2%
EZH2 0%
K27
KDM6A 1.3%
KDM6B 0.7%
EHMT1 0.4%
EHMT2 1.1%
SETDB1 1.1%
SUV39H2 0.2%
K9
KDM3A 0.4%
KDM3B 1.5%
KDM4A 0.2%
KDM4B 0.7%
KDM4C 0.7%
KDM4D 0.7%
MLL 1.7%
MLL2 1.3%
MLL3 1.5%
MLL4 1.5%
MLL5 1.1%
SETD1B 0.2%
SMYD1 0.2%
SMYD3 0.2%
K4
KDM5A 0.7%
KDM5B 0.9%
KDM5C 0.4%
KDM5D 0.4%
KDM1B 0.4%

Figure 4

A


Progression−free Survival (PFS)
Logrank P=0.695
Proportion of Patients
KDM6A mut/del (n=15)
KDM6A wild type (n=448)
Time from randomisation (months)
Overall Survival (OS)
Logrank P=0.0498
Proportion of Patients
KDM6A mut/del (n=15)
KDM6A wild type (n=448)
Time from randomisation (months)

B

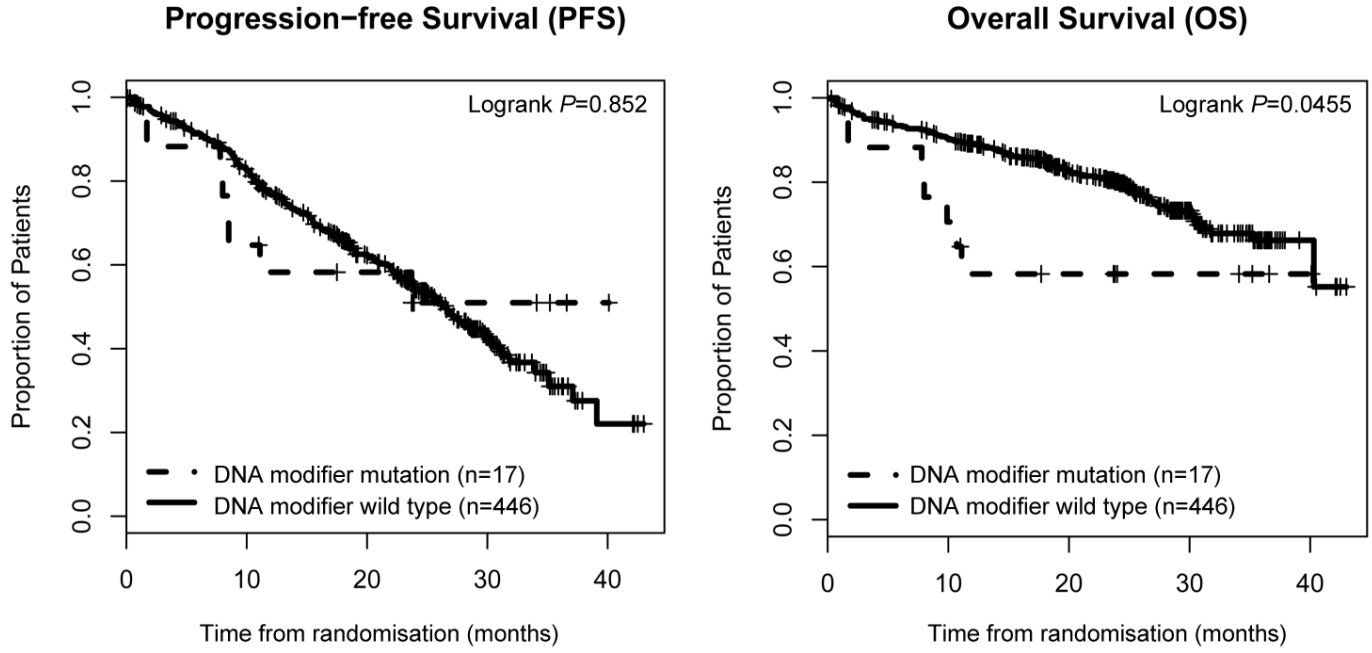
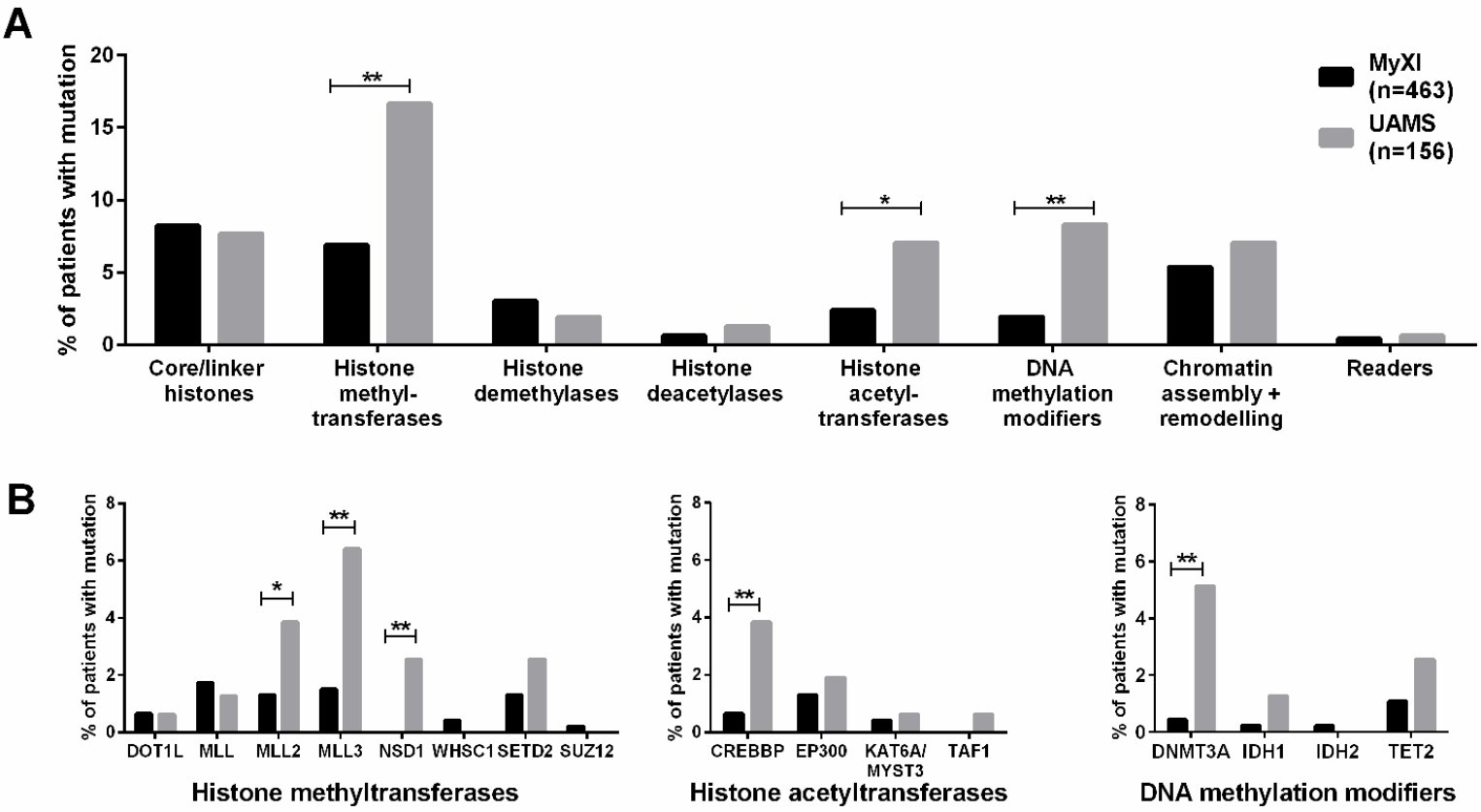

Progression−free Survival (PFS)
Logrank P=0.852
Proportion of Patients
DNA modifier mutation (n=17)
DNA modifier wild type (n=446)
Time from randomisation (months)
Overall Survival (OS)
Logrank P=0.0455
Proportion of Patients
DNA modifier mutation (n=17)
DNA modifier wild type (n=446)
Time from randomisation (months)

Figure 5


A
% of patients with mutation
20
15
10
5
0
MyXI (n=463)
UAMS (n=156)
**
*
**
Core/linker histones
Histone methyl-transferases
Histone demethylases
Histone deacetylases
Histone acetyl-transferases
DNA methylation modifiers
Chromatin assembly + remodelling
Readers
B
% of patients with mutation
8
6
4
2
0
*
**
**
DOT1L
MLL
MLL2
MLL3
NSD1
WHSC1
SETD2
SUZ12
Histone methyltransferases
**
CREBBP
EP300
KAT6A/ MYST3
TAF1
Histone acetyltransferases
**
DNMT3A
IDH1
IDH2
TET2
DNA methylation modifiers