Published in final edited form as:
Nat Genet. 2013 October ; 45(10): 1221–1225. doi:10.1038/ng.2733.

Common variation at 3q26.2, 6p21.33, 17p11.2 and 22q13.1 influences multiple myeloma risk

Daniel Chubb[#1], **Niels Weinhold**[#2], **Peter Broderick**[1], **Bowang Chen**[3], **David C Johnson**[4], **Asta Försti**[3,5], **Jayaram Vijayakrishnan**[1], **Gabriele Migliorini**[1], **Sara E Dobbins**[1], **Amy Holroyd**[1], **Dirk Hose**[2,6], **Brian A Walker**[4], **Faith E Davies**[4], **Walter A Gregory**[7], **Graham H Jackson**[8], **Julie A Irving**[9], **Guy Pratt**[10], **Chris Fegan**[11], **James AL Fenton**[12], **Kai Neben**[2], **Per Hoffmann**[13,14], **Markus M Nöthen**[13,14,15], **Thomas W Mühleisen**[13,14], **Lewin Eisele**[16], **Fiona M Ross**[17], **Christian Straka**[18], **Hermann Einsele**[19], **Christian Langer**[20], **Elisabeth Dörner**[2], **James M Allan**[9], **Anna Jauch**[21], **Gareth J Morgan**[4,**], **Kari Hemminki**[3,5,**], **Richard S Houlston**[1,**,¥], and **Hartmut Goldschmidt**[2,6,**]

[1]Division of Genetics and Epidemiology, Institute of Cancer Research, Surrey, UK

[2]Department of Internal Medicine V, University of Heidelberg, Heidelberg, Germany

¥Corresponding author: Richard S Houlston; richard.houlston@icr.ac.uk; Tel: +44 (0) 208 722 4175; Fax: +44 (0) 208 722 4365.
**These authors should be considered to have equal PI status

URLs

The R suite can be found at http://www.r-project.org/
Detailed information on the tag SNP panel can be found at http://www.illumina.com/
dbSNP: http://www.ncbi.nlm.nih.gov/projects/SNP/
HapMap: http://www.hapmap.org/
1000Genomes: http://www.1000genomes.org/
KBioscience: http://kbioscience.co.uk/
SNAP http://www.broadinstitute.org/mpg/snap/
IMPUTE: https://mathgen.stats.ox.ac.uk/impute/impute.html
EIGENSTRAT: http://genepath.med.harvard.edu/~reich/Software.htm
Wellcome Trust Case Control Consortium: www.wtccc.org.uk
METAL: www.sph.umich.edu/csg/abecasis/metal
Mendelian Inheritance In Man: http://www.ncbi.nlm.nih.gov/omim
Medical Research Council (MRC) Myeloma-IX trial: http://public.ukcrn.org.uk
Medical Research Council (MRC) Myeloma-XI trial: http://ctru.leeds.ac.uk/myelomaXI
USCS genome browser: http://genome.ucsc.edu/
HLA*IMP: https://oxfordhla.well.ox.ac.uk/hla/

Author Contributions

RSH and KH designed the study. RSH and GM obtained financial support in the UK and KH and HG in Germany. RSH drafted the manuscript. DC, BC and NW performed principal statistical and bioinformatic analyses; SED and GM performed additional statistical and bioinformatic analyses; PB coordinated UK laboratory analyses; JV and AH performed genotyping in the UK. DCJ managed and prepared Myeloma IX and Myeloma XI Case Study DNA samples. JMA conceived of the Newcastle-based Myeloma study (NMS). JMA established the study and supervised data collation and sample management of the NMS. JAI, GHJ, GP, JALF and CF developed protocols for recruitment of individuals with Myeloma and performed sample collection of cases within the NMS. HG, DH, KN and NW coordinated and managed German DNA samples, KH and AF coordinated German genotyping. HE, CL and CS ascertained and collected DSMM and Ulm Case Study samples and CL prepared DNA samples. ED and NW performed genotyping of German replication cases and controls. BAW performed UK expression analyses. FMR performed UK and AJ German FISH analyses. GJM, FED, WAG, GHJ and JAI performed ascertainment and collection of Case Study samples. PH, TWM and MMN performed and coordinated GWAS of German cases and controls. LE ascertained and managed the HNR sample. All authors contributed to the final paper.

Competing Financial Interests Statement

The authors declare no competing financial interests.

Europe PMC Funders Author Manuscripts

[3]German Cancer Research Center, Heidelberg, Germany

[4]Haemato-Oncology, Division of Pathology, Institute of Cancer Research, Surrey, UK

[5]Center for Primary Health Care Research, Lund University, Malmo, Sweden

[6]National Centre of Tumour Diseases, Heidelberg, Germany

[7]University of Leeds, Leeds, UK

[8]Royal Victoria Infirmary, Newcastle upon Tyne, UK

[9]Newcastle Cancer Centre, Northern Institute for Cancer Research, Medical School, Newcastle University, Newcastle-upon-Tyne, UK

[10]Department of Haematology, Birmingham Heartlands Hospital, Birmingham, UK

[11]Department of Haematology, School of Medicine, Cardiff University, Cardiff, UK

[12]Leeds Teaching Hospitals NHS Trust, Leeds, UK

[13]Institute of Human Genetics, University of Bonn, Germany

[14]Department of Genomics, University of Bonn, Germany

[15]German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany

[16]Institute for Medical Informatics, Biometry and Epidemiology, University Hospital Essen, University Duisburg-Essen, Germany

[17]Cytogenetics Group, Wessex Regional Cytogenetic Laboratory, Salisbury, UK

[18]Schön Klinik Starnberger See, Berg, Germany

[19]University Clinic of Würzburg, Würzburg, Germany

[20]Department of Internal Medicine III, University of Ulm, Ulm, Germany

[21]Institute of Human Genetics, University of Heidelberg, Germany

[#] These authors contributed equally to this work.

Abstract

To identify variants for multiple myeloma risk, we conducted a genome-wide association study with validation in additional series totaling 4,692 cases and 10,990 controls. We identified four risk loci at 3q26.2 (rs10936599, P=8.70x10^{-14}), 6p21.33 (rs2285803, *PSORS1C2;* P= 9.67x10^{-11}), 17p11.2 (rs4273077, *TNFRSF13B*; P=7.67x10^{-9}) and 22q13.1 (rs877529, *CBX7*; P=7.63x10^{-16}). These data provide further evidence for genetic susceptibility to this B-cell hematological malignancy and insight into the biological basis of predisposition.

Multiple myeloma (MM) is a malignancy of plasma cells1. Each year in the United States there are around 20,000 new cases of MM, and just over a half of that number die of the disease2. We have previously reported results of a genome-wide association study (GWAS) of MM based on an analysis of UK and German series and through fast track analysis of SNPs with the smallest P-values, identified risk loci at 2p23.3, 3p22.1 and 7p15.33. We have

subsequently conducted further follow-up analyses, making use of an expanded German GWAS, and identified four new susceptibility loci for MM.

The German GWAS data set previously reported4, after QC comprised 1,014 MM cases recruited through Heidelberg University genotyped using Illumina OmniExpress BeadChips. Genotype frequencies were compared with genotype data generated by the Heinz-Nixdorf Recall (HNR) study of 2,107 individuals5 from the German population who had been genotyped using Illumina Human Omni1-Quad BeadChips and Illumina OmniExpress BeadChips (Online Methods).

The UK GWAS previously reported3, after QC comprised 1,321 MM cases recruited from the UK Medical Research Council (MRC) Myeloma-IX trial6 genotyped using Illumina OmniExpress BeadChips. Genotype frequencies were compared with publicly accessible genotype data generated by the UK Wellcome Trust Case Control Consortium 2 (WTCCC2) study of 2,698 individuals from the 1958 British Birth Cohort (known as 58C)7 and 2,501 individuals from the UK Blood Service (UKBS) collections that had been genotyped using Illumina Human 1.2M-Duo Custom_v1 Array BeadChips (Online Methods).

Genotype data from the GWAS were filtered on the basis of pre-specified quality-control measures (Online Methods). Individual SNPs were excluded from further analysis if they showed deviation from the Hardy–Weinberg equilibrium with a $P<1.0x10^{-6}$ in controls, an individual SNP genotype yield <95%, or a minor allele frequency <1%. After filtering, 414,804 autosomal SNPs common to both case-control series were analyzable (Online Methods; Supplementary Figure 1; Supplementary Figure 2).

Prior to undertaking meta-analysis of the two GWAS, we searched for potential errors and biases in the datasets. Quantile-quantile plots of the genome-wide chi-squared values showed there was minimal inflation of the test statistics rendering substantial cryptic population substructure or differential genotype calling between cases and controls unlikely in either GWAS (genomic control inflation factor8, λ_{gc}=1.033 and 1.17 in UK and German GWAS, respectively; Supplementary Figure 3). For completeness principal components analysis was performed using the Eigenstrat9 software to determine the effects of population substructure on our findings ($\lambda_{corrected}$=1.014 and 1.029 in UK and German GWAS, respectively; Supplementary Figure 3).

Using data on all cases and controls from both GWAS, we derived joint odds ratios (ORs) and confidence intervals (CIs) under a fixed effects model for each SNP and associated *P*-values10. In the combined analysis we identified nine SNPs showing good evidence of association ($P<5.0\times10^{-6}$) and mapping to distinct loci not previously associated with MM risk (Supplementary Table 1). The *P*-value threshold used does not exclude the possibility that other SNPs represent genuine association signals but was simply a pragmatic strategy for prioritizing replication.

To validate our findings, we conducted a replication study of the nine SNPs, genotyping samples from three additional series: UK-replication-1, 812 MM cases ascertained through the UK MRC Myeloma-IX and XI trials and 1,110 controls; UK-replication-2, 396 MM cases collected through UK haematology centers and 992 controls; German-replication

1,149 MM cases collected through the German Myeloma Study Group (DSMM), Heidelberg University Clinic and Ulm University Clinic, and 1,582 regional controls (Online Methods). In the combined analysis, four SNPs rs10936599 (3q26.2; P=8.70x10^{-14}), rs2285803 (6p21.33; P= 9.67x10^{-11}), rs4273077 (17p11.2; P=7.67x10^{-9}) and rs877529 (22q13.1; P=7.63x10^{-16}) showed evidence for an association with MM which was genome-wide significant (Table 1, Online Methods, Supplementary Table 2).

rs10936599 at 3q26.2 (P=8.70x10^{-14}; Table 1) is responsible for the H717H polymorphism in the myoneurin gene (*MYNN*; MIM 606042). The rs10936599 G risk allele has previously also been shown to influence colorectal cancer risk11. While *MYNN* encodes a zinc finger protein of unknown function expressed principally in muscle rs10936599 (169,492,101bps) however maps within a 250Kb region of LD which also encompasses the telomerase RNA component gene (*TERC*; MIM 602322). Telomerase reactivation and telomerase-mediated elongation of shorter telomeres is a feature of MM12. Since carrier status for the rs10936599 G risk allele is associated with significantly longer telomeres12 *TERC* represents an attractive candidate for MM susceptibility. Moreover imputation of untyped genotypes in cases and controls using 1000 genomes data provided for a marginally stronger association at 3q26.2 with A allele of rs2293607, which maps 63bps 5' to *TERC* (P=6.2x10^{-10} compared with 1.3x10^{-9}; for rs10936599 in meta-analysis of GWAS data; Figure 1). The A allele of rs2293607 has recently been shown to be associated with *TERC* mRNA expression and longer telomeres *in vitro*12 supporting variation in *TERC* as the basis of the 3q26.2 cancer association.

rs2285803 (P=9.67x10^{-11}; Table 1; Figure 1) localizes in intron 5 of the putative psoriasis susceptibility gene, *PSORS1C1* (MIM 613525) at 6p21.33 (31,107,245bps). The 163 kb region of LD also encompasses *CCHCR1* (MIM 605310), *CDSN* (MIM 602593), transcription factor 19 (*TCF19*; MIM 600912) and POU domain, class 5, transcription factor 1 (*POU5F1*, MIM 164177) genes. While there is currently no evidence for *POU5F1* playing a role in MM intriguingly the gene encodes OCT3/OCT4 which regulates pluripotency, lineage commitment and regulates tissue-specific gene expression. Variation at 6p21.33 has previously been shown to be associated with follicular lymphoma (FL) and Hodgkin lymphoma (HL) risk. The associations for FL defined by rs6457327 in the HLA class I region13 and rs10484561 and rs2647012 in the HLA class II region14–15. The HL association at 6p21.33 is marked by rs6903608 in the HLA class II region16. The risk of MM associated with each of these SNPs was non-significant (Supplementary Table 3). To further investigate the rs2285803 signal for MM we imputed classical HLA alleles from SNP data from both GWASs using HLA*IMP17–18. The strongest HLA association was provided by HLA-DRB5*01 (P=1.42x10^{-5}; Supplementary Table 4) which was significantly weaker than provided by rs2285803 (P=3.07x10^{-8}). To evaluate the independence of associations, we conducted regression, jointly on rs2285803 and the imputed HLA alleles. Conditional analysis showed that most, but not all, of the MHC variation defined by SNP genotype could be explained for by rs2285803 (Supplementary Table 4).

rs4273077 (P=7.67x10^{-9}; Table 1) maps within intron 2 of the gene for *Homo sapiens* tumor necrosis factor receptor superfamily member 13B (*TNFRSF13B*; MIM 604907 at 17p11.2 (16,849,139bps; Figure 1). *TNFRSF13B* (alias *TACI*), represents a strong candidate for MM

predisposition *a priori*. TNFRSF13B is a key regulator of B and T-cell function being required for the development of transitional (T2) and mature-B lymphocytes, and regulation of normal B-cell homeostasis19. Variation at *TNFRSF13B* influences circulating IgG levels20 and Tnfrsf13b-/- mice show an expanded B-cell population with lymphoproliferation and lymphoma risk21. Since *TNFRSF13B* mutation is a risk factor for antibody-deficient (MIM 240500) and selective Ig deficiency (MIM 609529) associated with lymphoproliferation it is likely that loss of *TNFRSF13B* function impairs isotype switching. Primary MM cells with a high TNFRSF13B expression ($TACI^{high}$) resemble bone marrow plasma cells which depend on the interaction with the bone marrow environment. In contrast MM cells with a low expression of *TNFRSF13B* ($TACI^{low}$) resemble plasmablasts22. TACI-Ig, a soluble receptor blocking the TNFRSF13B ligands BAF and APRIL, inhibits the growth of $TACI^{high}$ but not $TACI^{low}$ myeloma cells in the SCID-hu model23.

rs877529 localizes to intron 2 of the gene encoding chromobox homolog 7 (*CBX7*; MIM 608457) at 22q13.1 (39,542,292bps; $P=7.63x10^{-16}$; Table 1; Figure 1). *CBX7* encodes a polycomb group protein. These proteins form part of a gene regulatory mechanism that determines cell fate during development as well as contributing to the control of normal cell growth and differentiation24. CBX7-mediated repression of transcription acts through Ink4a/Arf25, cooperating with Myc to promote aggressive B-cell lymphomagenesis with high levels of CBX7 being a feature of germinal center-derived follicular lymphoma26.

To explore whether any of the associations reflect *cis*-acting regulatory effects we studied mRNA expression in CD138-selected plasma cells27 and lymphoblastoid cell lines (LCLs)28–30 (Online Methods; Supplementary Table 5). Although we found no association between genotype and expression of either mRNA transcript, steady-state levels of RNA at a single time point may not adequately capture the impact of differential expression in tumourigenesis. To explore epigenetic profile of association signals we made use of chromatin state segmentation in lymphoblastoid cell lines data generated by the Encode Project31. rs2293607 maps to a region of active chromatin predicted by ENCODE data to be an active promoter and rs877529 maps within a strong enhancer element within *CBX7* (Figure 1).

Hierarchically MM can be broadly divided into hyperdiploid and non-hyperdiploid subtypes32–33. The latter is primarily composed of patients harboring IGH translocations, principally t(11;14)(q13;q32) and t(4;14)(p16;q32)34,35. Case-only analysis provided no evidence for a subtype specific association with genotype for rs10936599, rs2285803 or rs4237077 consistent with each variant having a generic effect on MM risk (Supplementary Table 6). In contrast rs877529 showed evidence, significant after correction for multiple testing, that the association is driven by non t(11;14) MM ($P=8.0x10^{-4}$; $P_{adj}= 0.016$).

Our findings provide further evidence for inherited genetic susceptibility to MM and insight into the development of this hematological malignancy. We estimate that the seven loci we have so far identified account for ~13% of the familial risk of MM. While the power of our study to detect the major common loci conferring risks of ≥1.3 was high we had low power to detect alleles with smaller effects and/or minor allele frequencies (MAFs) <0.1. By implication, variants with such profiles are likely to represent a much larger class of

susceptibility loci for MM, because of truly small effect sizes or submaximal LD with tagging SNPs. Thus, it is likely that a large number of variants remain to be discovered. This assertion is supported by the continued excess of associations observed over those expected, in addition to the regions studied herein. Further efforts to expand the scale of GWAS, in terms of both sample size and SNP coverage, and to increase the number of SNPs taken forward to large-scale replication may therefore identify additional risk variants. Finally as we have recently shown stratified analysis of MM by karyotype may lead additional subtype-specific risk variants4.

Online Methods

Ethics

Collection of samples and clinico-pathological information from subjects was undertaken with informed consent and relevant ethical review board approval in accordance with the tenets of the Declaration of Helsinki.

Genome-wide association study

UK-GWAS: Details of this study have been previously reported3. Briefly, 1,371 MM (ICD-10 C90.0; 469 male; mean age at diagnosis 63.9 years, SD 9.9) were ascertained through the UK Medical Research Council (MRC) Myeloma-IX trial6. Genotyping of cases was performed using Illumina Human OmniExpress-12 v1.0 arrays according to the manufacturer's protocols (Illumina, San Diego, USA). For controls, we used publicly accessible data generated by the Wellcome Trust Case Control Consortium from the 1958 Birth Cohort (58C; also known as the National Child Development Study)7 and National Blood Service (NBS). Genotyping of controls was conducted using Illumina Human 1-2M-Duo Custon_v1 Array chips. SNP calling was performed using Illuminus Software. Full details of genotyping, SNP calling and QC have been previously reported (www.wtccc.org.uk).

German-GWAS: The German-GWAS comprised 384 MM cases (229 male, mean age at diagnosis 54.5 years, SD 8.0) which were the subject of a previous publication3 and an additional series of 698 MM cases (389 male, mean age at diagnosis of 59 years; SD 9.3) recruited by the German Multiple Myeloma Study Group (GMMG), coordinated by the University Clinic, Heidelberg. All cases were genotyped using Illumina Human OmniExpress-12 v1.0 arrays according to the manufacturer's protocols (Illumina, San Diego, USA). For controls, we used genotype data on 2,132 healthy individuals, enrolled into the Heinz Nixdorf Recall (HNR) study5; of these 704 were genotyped using Illumina HumanOmni1-Quad_v1 and 1428 OmniExpress-12 v1.0.

Quality control of GWAS datasets

DNA samples with GenCall scores <0.25 at any locus were considered “no calls”. A SNP was deemed to have failed if <95% of DNA samples generated a genotype at the locus. Cluster plots were manually inspected for all SNPs considered for replication. The same quality control metrics on the new German GWAS data were applied as in our previous MM study3. We restricted analyses to samples for whom >95% of SNPs were successfully

genotyped, thus eliminating 10 samples (Supplementary Figure 1). We computed identity-by-state (IBS) probabilities for all pairs (cases and controls) to search for duplicates and closely related individuals amongst samples (defined as IBS ≥0.80, thereby excluding first-degree relatives). For all identical pairs the sample having the highest call rate was retained, eliminating 13 samples. To identify individuals who might have non-Western European ancestry, we merged our case and control data with phase II HapMap samples (60 western European [CEU], 60 Nigerian [YRI], 90 Japanese [JPT] and 90 Han Chinese [CHB]). For each pair of individuals we calculated genome-wide IBS distances on markers shared between HapMap and our SNP panel, and used these as dissimilarity measures upon which to perform principal component analysis. The first two principal components for each individual were plotted and any individual not present in the main CEU cluster was excluded from analyses. We removed 70 samples of non-CEU ancestry (some of which had poor call rates). We filtered out SNPs having a minor allele frequency [MAF] <1%, and a call rate <95% in cases or controls. We also excluded SNPs showing departure from Hardy-Weinberg equilibrium (HWE) at $P<10^{-6}$ in controls. For replication and validation analysis call rates were >95% per 384-well plate for each SNP; cluster plots were visually examined by two researchers.

Replication series and genotyping

UK-replication-1 comprised 812 MM cases (412 male) collected through the UK Medical Research Council (MRC) Myeloma-IX (n=95) and XI trials (n=717). Controls comprised 1,110 healthy individuals with self reported European ancestry (420 male, aged 18-69 years) with no personal history of malignancy ascertained through GEnetic Lung CAncer Predisposition Study (GELCAPS; n=536)36 and National Study of Colorectal Cancer Genetics (NSCCG; n=574)37 studies. All cases and controls were UK subjects.

UK-replication-2 comprised 396 MM cases (181 male; mean age at diagnosis 66.0 years, SD 12.5) collected through UK haematology departments (2001-present) including the Royal Marsden Hospitals NHS Trust (RMH). Controls were 992 healthy individuals (421 male, mean age 57.4 years, SD 12.3) with no personal history of malignancy who were the spouses of cancer patients ascertained by the ICR between 2000 and 2008.

German-replication comprised 1,149 cases collected by the German Myeloma Study Group (DSMM), GMMG, University Clinic, Heidelberg and University Clinic, Ulm (676 males; mean age at diagnosis 57.6 years, SD 9.8). The control population was composed of 1,582 healthy German blood donors who were recruited between 2004 and 2007 by the Institute of Transfusion Medicine and Immunology, University of Mannheim, Germany (885 male, mean age 55.8 years, SD 10.0).

Replication genotyping was performed using competitive allele-specific PCR KASPar chemistry (KBiosciences Ltd, Hertfordshire, UK). All primers and probes used are available on request. Samples having SNP call rates of <90% were excluded from the analysis. To ensure quality of genotyping in all assays, at least two negative controls and 1-2% duplicates (showing a concordance >99.99%) were genotyped. To exclude technical artifact in genotyping we performed cross-platform validation of 384 samples and sequenced a set of

384 randomly selected samples from each case and control series to confirm genotyping accuracy (concordance >99.9%).

Sample preparation

For German cases DNA was prepared from EDTA-venous blood samples, 100% of the original GWAS, 42% of the additional GWAS, and 91% of the replication samples; for the remaining cases, the source was the CD138-negative fraction of bone marrow cells, with < 5% contamination by tumor cells. For all UK cases DNA was prepared from EDTA-venous blood samples. Samples were obtained prior to delivery of chemotherapy in the vast majority of UK cases and at least 80% of the German cases. All DNAs were extracted using Qiagen FlexiGene or QIAamp methodologies and quantified using PicoGreen (Invitrogen).

Statistical and bioinformatic analysis

Main analyses were undertaken using R (v2.6), Stata v.10 (State College, Texas, US) and PLINK (v1.06)38 software. Odds ratios (ORs) and associated 95% confidence intervals (CIs) along with associated *P*-values were calculated by unconditional logistic regression. The adequacy of the case-control matching and possibility of differential genotyping of cases and controls were formally evaluated using quantile-quantile (Q-Q) plots of test statistics. The inflation factor λ was based on the 90% least significant SNPs8. We undertook adjustment for possible population substructure using Eigenstrat software. Meta-analysis was conducted using standard methods10. Cochran's Q statistic to test for heterogeneity10 and the I^2 statistic to quantify the proportion of the total variation due to heterogeneity were calculated39. I^2 values ≥75% are considered characteristic of large heterogeneity39–40. To conduct a pooled analysis incorporating Eigenstrat adjusted *P*-values from the GWAS we used the weighted Z-method implemented in the program METAL41. Since not all the HNR controls were genotyped using the same Illumina array the robustness of genomewide associations was formally assessed by deriving ORs for the different German case-control combinations and incorporation of these data in meta-analysis (Supplementary Figure 4). We examined each SNP for dose response by comparing 1-d.f. and 2-d.f. logistic regression models, adjusting for stage using a likelihood ratio test, and examined the combined effects of multiple SNPs by evaluating the effect of adding an interaction term on the model by using a likelihood ratio test and adjusting for stage. Associations by tumor karyotype were examined by logistic regression in case-only analyses. The sibling relative risk attributable to a given SNP was calculated using the formula:

$$\lambda^{\star}=\frac{p(pr_2+qr_1)^2+q(pr_1+q)^2}{[p^2r_2+2pqr_1+q^2]^2}$$

where p is the population frequency of the minor allele, $q=1-p$, and r_1 and r_2 are the relative risks (estimated as OR) for heterozygotes and rare homozygotes, relative to common homozygotes. Assuming a multiplicative interaction, the proportion of the familial risk attributable to a SNP was calculated as $\log\lambda^*/\log\lambda_0$, where λ_0 is the overall familial relative risk estimated from epidemiological studies, assumed to be 2.4542.

Europe PMC Funders Author Manuscripts

Prediction of the untyped SNPs was carried out using IMPUTEv2, based on the 1000 genomes phase 1 integrated variant set (b37) from March 2012. Imputed data were analysed using SNPTEST v2 to account for uncertainties in SNP prediction and meta-analysis was performed using METAv1.443. LD metrics were calculated in plink using 1000 genomes data and plotted using SNAP. LD blocks were defined on the basis of HapMap recombination rate (cM/Mb) as defined using the Oxford recombination hotspots44 and on the basis of distribution of confidence intervals defined by Gabriel *et al*45. We imputed classical HLA alleles from GWAS SNPs using HLA*IMP17–18

To explore epigenetic profile of association signals we made use of chromatin state segmentation in lymphoblastoid cell lines data generated by the Encode Project31. The states were inferred from ENCODE Histone Modification data (H4K20me1, H3K9ac, H3K4me3, H3K4me2, H3K4me1, H3K36me3, H3K27me3, H3K27ac and CTCF) binarized using a multivariate Hidden Markov Model.

Karotyping and Fluorescence *in situ* hybridization (FISH)

Conventional cytogenetic studies of multiple myeloma cells were conducted using standard karotyping methodologies, and standard criteria for the definition of a clone were applied. FISH and ploidy classification of UK samples was conducted using the methodology described by Chiecchio *et al 46*. FISH and ploidy classification of German samples was performed as previously described47. The XL IGH Break Apart probe (MetaSystems, Altlussheim Germany) was used to detect any IGH translocation in German samples.

Relationship between SNP genotype and mRNA expression

To examine for a relationship between SNP genotype and mRNA expression in MM we made use of Affymetrix Human Genome U133+2.0 array data on the plasma cells from 192 MM patients from the MRC Myeloma IX trial27. To assay *TERC* which was not captured on the U133+2.0 array we made use of Affymetrix GeneChip miRNA 2.0 data. To examine for a relationship between SNP genotype and expression levels in lymphocytes we made use of publicly available expression data generated on lymphoblastoid cell lines from HapMap329 Geneva GenCord individuals30 and the MuTHER resource28 using Sentrix Human-6 Expression BeadChips (Illumina, San Diego, USA)48–49.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Leukaemia Lymphoma Research and Myeloma UK provided principal funding for this study in the UK. Additional funding was provided by Cancer Research UK (C1298/A8362 supported by the Bobby Moore Fund) and the NHS via the Biological Research Centre of the National Institute for Health Research at the Royal Marsden Hospital NHS Trust. This study made use of genotyping data on the 1958 Birth Cohort. Genotyping data on controls was generated by the Wellcome Trust Sanger Institute. A full list of the investigators who contributed to the generation of the data are available from http://www.wtccc.org.uk. In Germany (Heidelberg) funding was provided to Dietmar-Hopp-Stiftung Walldorf, the University Hospital Heidelberg, Deutsche Krebshilfe and the Systems Medicine funding of the German Ministry of Education and Science. We are grateful to all investigators who contributed to NSCCG and GELCAPS, from which controls in the replication were drawn. The GWAS made use of genotyping data from the population based HNR study. The HNR study is supported by the Heinz Nixdorf Foundation

(Germany). Additionally, the study is funded by the German Ministry of Education and Science and the German Research Council (DFG; Project SI 236/8-1, SI236/9-1, ER 155/6-1, and the DFG CRU 216). Funding was provided to LE by the Medical Faculty of the University Hospital of Essen (IFORES). The genotyping of the Illumina HumanOmni-1 Quad BeadChips of the HNR subjects was financed by the German Center for Neurodegenerative Disorders (DZNE), Bonn. We are extremely grateful to all investigators who contributed to the generation of this dataset. The German replication controls were collected by Peter Bugert, Institute of Transfusion Medicine and Immunology, Medical Faculty Mannheim, Heidelberg University, German Red Cross Blood Service of Baden-Württemberg-Hessen, Mannheim, Germany. Finally, we are grateful to all the patients and investigators at the individual centers for their participation. We also thank the staff of the CTRU University of Leeds and the NCRI haematology Clinical Studies Group.

References

1. Kyle RA, Rajkumar SV. Multiple myeloma. N Engl J Med. 2004; 351:1860–73. [PubMed: 15509819]
2. Palumbo A, Anderson K. Multiple myeloma. N Engl J Med. 2011; 364:1046–60. [PubMed: 21410373]
3. Broderick P, et al. Common variation at 3p22.1 and 7p15.3 influences multiple myeloma risk. Nat Genet. 2012; 44:58–61.
4. Weinhold N, et al. The CCND1 G870A polymorphism is a risk factor for t(11;14)(q13;q32) multiple myeloma. Nature Genetics. 2013 (in press) Nature Genetics (in press).
5. Schmermund A, et al. Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: rationale and design of the Heinz Nixdorf RECALL Study. Risk Factors, Evaluation of Coronary Calcium and Lifestyle. Am Heart J. 2002; 144:212–8. [PubMed: 12177636]
6. Morgan GJ, et al. First-line treatment with zoledronic acid as compared with clodronic acid in multiple myeloma (MRC Myeloma IX): a randomised controlled trial. Lancet. 2010; 376:1989–99. [PubMed: 21131037]
7. Power C, Elliott J. Cohort profile: 1958 British birth cohort (National Child Development Study). Int J Epidemiol. 2006; 35:34–41. [PubMed: 16155052]
8. Clayton DG, et al. Population structure, differential bias and genomic control in a large-scale, case-control association study. Nat Genet. 2005; 37:1243–6. [PubMed: 16228001]
9. Price AL, et al. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006; 38:904–9. [PubMed: 16862161]
10. Pettiti, D. Meta-analysis decision analysis and cost-effectivness analysis. Oxford University Press; 1994.
11. Houlston RS, et al. Meta-analysis of three genome-wide association studies identifies susceptibility loci for colorectal cancer at 1q41, 3q26.2, 12q13.13 and 20q13.33. Nat Genet. 2010; 42:973–7. [PubMed: 20972440]
12. Jones AM, et al. TERC polymorphisms are associated both with susceptibility to colorectal cancer and with longer telomeres. Gut. 2012; 61:248–54. [PubMed: 21708826]
13. Skibola CF, et al. Genetic variants at 6p21.33 are associated with susceptibility to follicular lymphoma. Nat Genet. 2009; 41:873–5. [PubMed: 19620980]
14. Conde L, et al. Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. Nat Genet. 2010; 42:661–4. [PubMed: 20639881]
15. Smedby KE, et al. GWAS of follicular lymphoma reveals allelic heterogeneity at 6p21.32 and suggests shared genetic susceptibility with diffuse large B-cell lymphoma. PLoS Genet. 2011; 7:e1001378. [PubMed: 21533074]
16. Enciso-Mora V, et al. A genome-wide association study of Hodgkin's lymphoma identifies new susceptibility loci at 2p16.1 (REL), 8q24.21 and 10p14 (GATA3). Nat Genet. 2010; 42:1126–30. [PubMed: 21037568]
17. Leslie S, Donnelly P, McVean G. A statistical method for predicting classical HLA alleles from SNP data. Am J Hum Genet. 2008; 82:48–56 (2008). [PubMed: 18179884]
18. Dilthey AT, Moutsianas L, Leslie S, McVean G. HLA*IMP--an integrated framework for imputing classical HLA alleles from SNP genotypes. Bioinformatics. 27:968–72.

19. Gross JA, et al. TACI-Ig neutralizes molecules critical for B cell development and autoimmune disease. impaired B cell maturation in mice lacking BLyS. Immunity. 2001; 15:289–302. [PubMed: 11520463]
20. Liao M, et al. Genome-wide association study identifies common variants at TNFRSF13B associated with IgG level in a healthy Chinese male population. Genes Immun. 2012
21. Seshasayee D, et al. Loss of TACI causes fatal lymphoproliferation and autoimmunity, establishing TACI as an inhibitory BLyS receptor. Immunity. 2003; 18:279–88. [PubMed: 12594954]
22. Moreaux J, et al. The level of TACI gene expression in myeloma cells is associated with a signature of microenvironment dependence versus a plasmablastic signature. Blood. 2005; 106:1021–30. [PubMed: 15827134]
23. Yaccoby S, et al. Atacicept (TACI-Ig) inhibits growth of TACIhigh primary myeloma cells in SCID-hu mice and in coculture with osteoclasts. Leukemia. 2007; 22:406–413. [PubMed: 18046446]
24. Gil J, Bernard D, Peters G. Role of polycomb group proteins in stem cell self-renewal and cancer. DNA Cell Biol. 2005; 24:117–25. [PubMed: 15699631]
25. Aguilo F, Zhou MM, Walsh MJ. Long noncoding RNA, polycomb, and the ghosts haunting INK4b-ARF-INK4a expression. Cancer Res. 71:5365–9.
26. Scott CL, et al. Role of the chromobox protein CBX7 in lymphomagenesis. Proc Natl Acad Sci U S A. 2007; 104:5389–94. [PubMed: 17374722]
27. Walker BA, et al. Integration of global SNP-based mapping and expression arrays reveals key regions, mechanisms, and genes important in the pathogenesis of multiple myeloma. Blood. 2006; 108:1733–43. [PubMed: 16705090]
28. Nica AC, et al. The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. PLoS Genet. 2011; 7:e1002003. [PubMed: 21304890]
29. Stranger BE, et al. Patterns of cis regulatory variation in diverse human populations. PLoS Genet. 2012; 8:e1002639. [PubMed: 22532805]
30. Dimas AS, et al. Common regulatory variation impacts gene expression in a cell type-dependent manner. Science. 2009; 325:1246–50. [PubMed: 19644074]
31. Ernst J, K M. Discovery and characterization of chromatin states for systematic annotation of the human genome. Nat Biotechnol. 2010; 28:817–825. [PubMed: 20657582]
32. Dewald GW, Kyle RA, Hicks GA, Greipp PR. The clinical significance of cytogenetic studies in 100 patients with multiple myeloma, plasma cell leukemia, or amyloidosis. Blood. 1985; 66:380–90. [PubMed: 3926026]
33. Debes-Marun CS, et al. Chromosome abnormalities clustering and its implications for pathogenesis and prognosis in myeloma. Leukemia. 2003; 17:427–36. [PubMed: 12592343]
34. Fonseca R, et al. International Myeloma Working Group molecular classification of multiple myeloma: spotlight review. Leukemia. 2009; 23:2210–21. [PubMed: 19798094]
35. Walker BA, et al. A compendium of myeloma-associated chromosomal copy number abnormalities and their prognostic value. Blood. 2010; 116:e56–65. [PubMed: 20616218]
36. Eisen T, Matakidou A, Houlston R. Identification of low penetrance alleles for lung cancer: the GEnetic Lung CAncer Predisposition Study (GELCAPS). BMC Cancer. 2008; 8:244. [PubMed: 18715499]
37. Penegar S, et al. National study of colorectal cancer genetics. Br J Cancer. 2007; 97:1305–9. [PubMed: 17895893]
38. Purcell S, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81:559–75. [PubMed: 17701901]
39. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002; 21:1539–58. [PubMed: 12111919]
40. Ioannidis JP, Ntzani EE, Trikalinos TA. 'Racial' differences in genetic effects for complex diseases. Nat Genet. 2004; 36:1312–8. [PubMed: 15543147]
41. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010; 26:2190–1. [PubMed: 20616382]

42. Altieri A, Chen B, Bermejo JL, Castro F, Hemminki K. Familial risks and temporal incidence trends of multiple myeloma. Eur J Cancer. 2006; 42:1661–70. [PubMed: 16753294]
43. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet. 2007; 39:906–13. [PubMed: 17572673]
44. Myers S, Bottolo L, Freeman C, McVean G, Donnelly P. A fine-scale map of recombination rates and hotspots across the human genome. Science. 2005; 310:321–4. [PubMed: 16224025]
45. Gabriel SB, et al. The structure of haplotype blocks in the human genome. Science. 2002; 296:2225–9. [PubMed: 12029063]
46. Chiecchio L, et al. Deletion of chromosome 13 detected by conventional cytogenetics is a critical prognostic factor in myeloma. Leukemia. 2006; 20:1610–7. [PubMed: 16826223]
47. Neben K, et al. Combining information regarding chromosomal aberrations t(4;14) and del(17p13) with the International Staging System classification allows stratification of myeloma patients undergoing autologous stem cell transplantation. Haematologica. 2010; 95:1150–7. [PubMed: 20220069]
48. Stranger BE, et al. Genome-wide associations of gene expression variation in humans. PLoS Genet. 2005; 1:e78. [PubMed: 16362079]
49. Stranger BE, et al. Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science. 2007; 315:848–53. [PubMed: 17289997]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

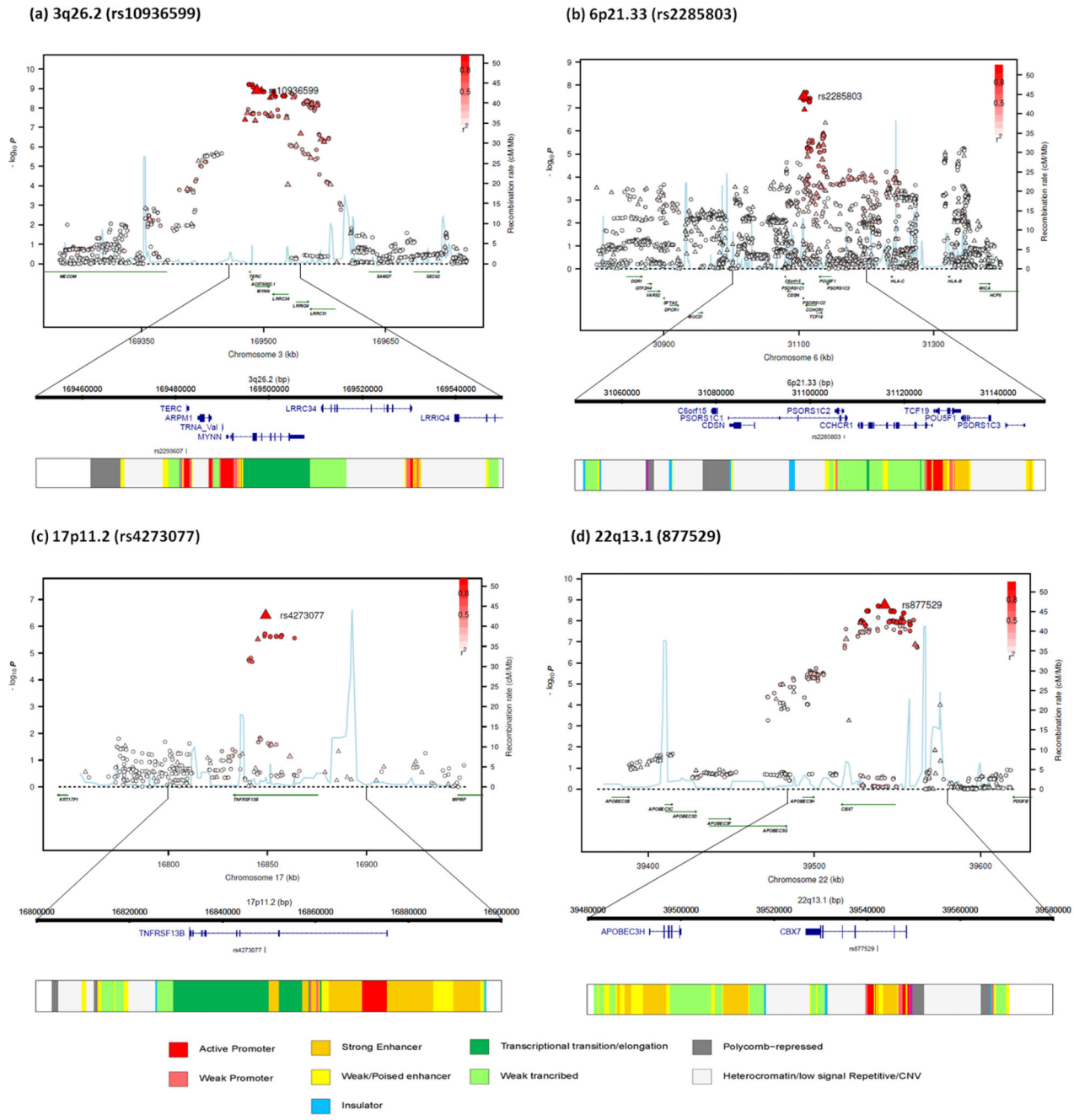

Figure 1. Regional plots of association results and recombination rates for the 3q26.2, 6p21.33, 17p11.2 and 22q13.1 susceptibility loci.

(a-d) Association results of both genotyped (triangles) and imputed (circles) SNPs in the GWAS samples and recombination rates for rates within the four loci: 3q26.2, 6p21.33, 17p11.2 and 22q13.1. For each plot, $-\log_{10}$ *P* values (*y* axis) of the SNPs are shown according to their chromosomal positions (*x* axis). The top genotyped SNP in each combined analysis is a large triangle and is labeled by its rsID. The color intensity of each symbol reflects the extent of LD with the top genotyped SNP: white (r^2=0) through to dark

red (r^2=1.0). Genetic recombination rates (cM/Mb), estimated using HapMap CEU samples, are shown with a light blue line. Physical positions are based on NCBI build 36 of the human genome. Also shown are the relative positions of genes and transcripts mapping to each region of association. Genes have been redrawn to show the relative positions; therefore, maps are not to physical scale. The lower panel shows the region of interest together with all transcripts and chromatin state segmentation track (ChromHMM) for lymphoblastoid cells using data from the HapMap Encode Project.

Europe PMC Funders Author Manuscripts

Table 1
Summary results for SNPs associated with multiple myeloma risk.

	Risk allele	RAF[a]	Case genotypes			RAF[a]	Control genotypes			OR[b]	95% CI[c]	*P*-value	$P_{adjusted}$[d]
rs10936599 (3q26.2)	**G**		**GG**	**AG**	**AA**		**GG**	**AG**	**AA**				
UK-GWAS		0.80	843	429	49	0.75	2960	1914	325	1.31	1.18-1.46	$4.33x10^{-7}$	$5.18x10^{-7}$
German-GWAS		0.79	632	332	49	0.75	1187	778	142	1.25	1.10-1.41	$6.62x10^{-4}$	$1.48x10^{-3}$
UK replication 1		0.80	520	259	30	0.76	628	415	63	1.32	1.13-1.55	$4.98x10^{-4}$	-
UK replication 2		0.79	244	126	18	0.75	559	372	56	1.23	1.01-1.51	$4.22x10^{-2}$	-
German replication		0.78	714	363	66	0.76	898	585	89	1.16	1.02-1.31	$2.56x10^{-2}$	-
Combined										**1.26**	**1.18-1.33**	**$8.70x10^{-14}$**	**$1.74x10^{-13}$**
												P_{het}=0.60, I^2=0%	
rs2285803 (6p21.3)	**A**		**AA**	**AG**	**GG**		**AA**	**AG**	**GG**				
UK-GWAS		0.32	125	603	593	0.28	444	2055	2699	1.21	1.10-1.32	$6.67x10^{-5}$	$7.64x10^{-5}$
German-GWAS		0.36	129	462	423	0.31	226	833	1047	1.24	1.11-1.39	$1.15x10^{-4}$	$1.18x10^{-4}$
UK replication 1		0.32	78	362	364	0.28	88	424	560	1.22	1.06-1.41	$5.11x10^{-3}$	-
UK replication 2		0.29	32	152	193	0.26	51	402	510	1.14	0.94-1.38	$1.82x10^{-1}$	-
German replication		0.33	130	491	521	0.30	140	674	752	1.12	1.00-1.26	$5.86x10^{-2}$	-
Combined										**1.19**	**1.13-1.26**	**$9.67x10^{-11}$**	**$1.18x10^{-10}$**
												P_{het}=0.70, I^2=0%	
rs4273077 (17p11.2)	**G**		**GG**	**AG**	**AA**		**GG**	**AG**	**AA**				
UK-GWAS		0.12	15	284	1022	0.10	48	926	4221	1.24	1.08-1.42	$1.88x10^{-3}$	$2.65x10^{-3}$
German-GWAS		0.14	25	239	750	0.11	27	390	1690	1.40	1.20-1.64	$2.80x10^{-5}$	$6.17x10^{-4}$
UK replication 1		0.12	18	148	629	0.09	12	179	915	1.28	1.04-1.57	$1.96x10^{-2}$	-
UK replication 2		0.11	3	77	304	0.10	8	178	805	1.12	0.85-1.48	$4.20x10^{-1}$	-
German replication		0.12	17	252	876	0.11	21	298	1244	1.17	0.99-1.38	$6.79x10^{-2}$	-
Combined										**1.26**	**1.16-1.36**	**$7.67x10^{-9}$**	**$1.41x10^{-7}$**
												P_{het}=0.50 I^2=0%	
rs877529 (22q13.1)	**A**		**AA**	**AG**	**GG**		**AA**	**AG**	**GG**				
UK-GWAS		0.51	346	654	321	0.44	1000	2560	1633	1.33	1.22-1.45	$1.08x10^{-10}$	$9.11x10^{-11}$
German-GWAS		0.45	214	483	317	0.43	389	1026	692	1.09	0.98-1.21	$1.18x10^{-1}$	$1.09x10^{-1}$
UK replication 1		0.49	176	436	192	0.44	195	555	322	1.24	1.08-1.41	$2.01x10^{-3}$	-
UK replication 2		0.47	86	192	109	0.42	166	485	327	1.24	1.05-1.47	$1.21x10^{-2}$	-
German replication		0.46	238	586	321	0.41	274	754	544	1.23	1.10-1.37	$2.69x10^{-4}$	-
Combined										**1.23**	**1.17-1.29**	**$7.63x10^{-16}$**	**$2.29x10^{-16}$**
												P_{het}=0.09, I^2=51%	

[a] Risk allele frequency (RAF).

[b] Odds ratio.

[c] 95% Confidence Interval.

[d] Eigenstrat adjusted *P*-values.