Genome-wide association study identifies multiple susceptibility loci for multiple myeloma

Jonathan S. Mitchell1,*, Ni Li1,*, Niets Weinhold2,3,*, Asta Försti4,5,*, Mina Ali6,*, Mark van Duin7,*, Gudmar Thorleifsson8, David C. Johnson9, Bowang Chen4, Britt-Marie Halvarsson6, Daniel F. Gudbjartsson8,10, Rowan Kuiper7, Owen W. Stephens2, Uta Bertsch3,11, Peter Broderick1, Chiara Campo4, Hermann Einsele12, Walter A. Gregory13, Urban Gullberg6, Marc Henrion1, Jens Hillengass3, Per Hoffmann14,15, Graham H. Jackson16, Ellinor Johnsson6, Magnus Jöud6,17, Sigurður Y. Kristinsson18, Sig Lennhoff19, Oleg Lenive1, Ulf-Henrik Mellqvist20, Gabriele Migliorini1, Hareth Nahi21, Sven Nelander22, Jolanta Nickel9, Markus M. Nöthen14,23, Thorunn Rafnar8, Fiona M. Ross24, Miguel Inacio da Silva Filho4, Bhairavi Swaminathan6, Hauke Thomsen4, Ingemar Turesson19, Annette Vangsted25, Ulla Vogel26, Anders Waage27, Brian A. Walker2, Anna-Karin Wihlborg6, Annemieke Broyle17, Faith E. Davies2, Unnur Thorsteinssdotter8,28, Christian Langer29, Markus Hansson6,19, Martin Kaiser9, Pieter Sonneveld7, Kari Stefansson8,21, Gareth J. Morgan2,22, Hartmut Goldschmidt3,11,21, Kari Hemminki4,5,14, Björn Nilsson6,17,30,30, & Richard S. Houlston1,22,3

Multiple myeloma (MM) is a plasma cell malignancy with a significant heritable basis. Genome-wide association studies have transformed our understanding of MM predisposition, but individual studies have had limited power to discover risk loci. Here we perform a meta-analysis of these GWAS, add a new GWAS and perform replication analyses resulting in 9,866 cases and 239,188 controls. We confirm all nine known risk loci and discover eight new loci at 6p22.3 (rs34229995, \(P = 1.31 \times 10^{-8}\)), 6q21 (rs9372120, \(P = 9.09 \times 10^{-15}\)), 7q36.1 (rs7781265, \(P = 9.71 \times 10^{-9}\)), 8q24.21 (rs1948915, \(P = 4.20 \times 10^{-11}\)), 9p23.1 (rs2811710, \(P = 1.72 \times 10^{-13}\)), 10p12.1 (rs2790457, \(P = 1.77 \times 10^{-8}\)), 16q23.1 (rs7193541, \(P = 5.00 \times 10^{-12}\)) and 20q13.13 (rs6066635, \(P = 1.36 \times 10^{-13}\)), which localize in or near to JARID2, ATG5, SMARCD3, CCAT1, CDKN2A, WAC, RFWD3 and PREX1. These findings provide additional support for a polygenic model of MM and insight into the biological basis of tumour development.

1 Division of Genetics and Epidemiology, The Institute of Cancer Research, 15 Colworth Road, Sutton, Surrey SM2 5NG, UK. 2 Myeloma Institute for Research and Therapy, University of Arkansas for Medical Sciences, Little Rock, Arkansas 72205, USA. 3 Department of Internal Medicine V, University of Heidelberg, 69117 Heidelberg, Germany. 4 German Cancer Research Center, 69120 Heidelberg, Germany. 5 Center for Primary Health Care Research, Lund University, SE-205 02 Malmo, Sweden. 6 Hematology and Transfusion Medicine, Department of Laboratory Medicine, BMC B13, SE-221 84 Lund, Sweden. 7 Hematology, Erasmus MC Cancer Institute, 3075 EA Rotterdam, The Netherlands. 8 deCODE Genetics, Sturlugata 8, IS-101 Reykjavik, Iceland. 9 Division of Molecular Pathology, The Institute of Cancer Research, Surrey SM2 5NG, UK. 10 School of Engineering and Natural Sciences, University of Iceland, IS-101 Reykjavik, Iceland. 11 National Centre of Tumor Diseases, 69120 Heidelberg, Germany. 12 University Clinic of Würzburg, 97080 Würzburg, Germany. 13 Clinical Trials Research Unit, University of Leeds, Leeds LS2 9HP, UK. 14 Institute of Human Genetics, University of Bonn, D-53127 Bonn, Germany. 15 Division of Medical Genetics, Department of Biomedicine, University of Basel, 4003 Basel, Switzerland. 16 Royal Victoria Infirmary, Newcastle upon Tyne, NE1 4LP, UK. 17 Clinical Immunology and Transfusion Medicine, Laboratory Medicine, Office of Medical Services, SE-221 85 Lund, Sweden. 18 Department of Hematology, Landspítali, National University Hospital of Iceland, IS-101 Reykjavik, Iceland. 19 Hematology Clinic, Skåne University Hospital, SE-221 85 Lund, Sweden. 20 Section of Hematology, Sahlgrenska University Hospital, Gothenburg 413 45, Sweden. 21 Center for Hematology and Regenerative Medicine, Karolinska Institutet, SE-171 77 Stockholm, Sweden. 22 Rudbeck Laboratory, Department of Immunology, Pathology and Genetics, Uppsala University, SE-751 08 Uppsala, Sweden. 23 Department of Genomics, Life & Brain Center, University of Bonn, D-53127 Bonn, Germany. 24 Wessex Regional Genetics Laboratory, University of Southampton, Salisbury SP2 8BJ, UK. 25 Department of Haematology, University Hospital of Copenhagen at Rigshospitalet, Blegdamsvej 9, DK-2100 Copenhagen, Denmark. 26 National Research Centre for the Working Environment, DK-2100 Copenhagen, Denmark. 27 Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology, Box 8805, N-7491 Trondheim, Norway. 28 Faculty of Medicine, University of Iceland, IS-101 Reykjavik, Iceland. 29 Department of Internal Medicine III, University of Ulm, D-89081 Ulm, Germany. 30 Broad Institute, 7 Cambridge Center, Cambridge, Massachusetts 02142, USA. * These authors jointly supervised the work. Correspondence and requests for materials should be addressed to K.H. (email: K.Hemminki@dkfz-heidelberg.de) or to B.N. (email: bjorn.nilsson@med.lu.se) or to R.S.H. (email: richard.houlston@icr.ac.uk).
Multiple myeloma (MM) is a malignancy of plasma cells that has a significant genetic component as evidenced by the two- to fourfold increased risk shown in relatives of MM patients. Our understanding of MM susceptibility has been transformed by recent genome-wide association studies (GWASs), which have identified the first risk alleles for MM and its precursor condition monoclonal gammopathy of unknown significance. Although projections indicate that additional risk variants for MM can be discovered by GWAS, the statistical power of these individual studies is limited.

To gain comprehensive insight into MM predisposition, we performed a meta-analysis of these GWAS, new GWAS and replication comprising 9,866 cases and 239,188 controls. We confirmed all nine known risk loci and discovered eight new risk loci for MM. Our findings provide further insights into the genetic and biological basis of MM predisposition.

**Results**

**Association analysis.** To identify new MM susceptibility loci, we analysed genome-wide association data from six populations of European ancestry (Supplementary Tables 1 and 2): a new sample set from the Netherlands, two previously reported sample sets from United Kingdom and Germany, to which we added additional cases, and three previously published sample sets from the Netherlands, United Kingdom, Germany, Sweden/Norway and Denmark sample series totalling 2,547 cases and 4,803 controls. Meta-analysing the discovery and replication samples, we identified genome-wide significant associations for MM with eight previously unreported loci (Table 1 and Supplementary Tables 8 and 9) at 6p22.3 (rs34229995, \( P = 1.31 \times 10^{-8} \)), 6q21 (rs9372120, \( P = 9.09 \times 10^{-15} \)), 7q36.1 (rs7781265, \( P = 9.71 \times 10^{-9} \)), 8q24.21 (rs1948915, \( P = 4.20 \times 10^{-11} \)), 9p21.3 (rs2811710, \( P = 1.72 \times 10^{-15} \)), 10p12.1 (rs2790457, \( P = 1.77 \times 10^{-8} \)), 16q21 (rs6066835, \( P = 1.36 \times 10^{-13} \)), 20q13.13 (rs7781265, \( P = 5.00 \times 10^{-12} \)) and 20q13.13 (rs6066835, \( P = 1.36 \times 10^{-13} \)). We also observed two promising associations (that is, \( P<5.0 \times 10^{-7} \)) at 6q27 (rs1034447) and at 7q22.3 (rs17507636) (Supplementary Tables 8 and 9). Conditional analysis of GWAS showed no evidence for additional independent signals at the loci.

The 6q21 association marked by rs9372120 (Fig. 2) maps to intron 6 of *ATG5* (*Homo sapiens* autophagy related 5). The 8q24.21 variant rs1948915 maps to *CCT1* (colon cancer-associated transcript 1; Fig. 2). The same region at 8q24.21 harbours multiple independent loci with different tumour specificities, including the B-cell malignancies diffuse B-cell lymphoma, Hodgkin’s lymphoma and chronic lymphocytic leukaemia. With the possible exception of chronic lymphocytic leukaemia, the linkage disequilibrium (LD) blocks defining these identified cancer risk loci are distinct from the 8q24.21 MM association signal (pairwise LD metrics \( r^2<0.03 \); Supplementary Table 10). The 9p21.3 variant rs2811710 maps to intron 1 of *CDKN2A/p16INK4A* (cyclin-dependent kinase inhibitor 2A, Fig. 2). Although the 9p21.3 region is a susceptibility locus for multiple tumour types including breast and lung cancer, glioma and acute lymphoblastic leukaemia, the 8q24.21 region is specific for MM (Supplementary Table 11). The 16q23.1 variant rs193541 (association localizes to intron 3 of the gene *CCAT1*; Fig. 2). The 10p12.1 (rs2790457) association localizes to intron 2 of *SMARCD3* (related, matrix-associated, actin-dependent regulator of chromatin, subfamily d, member 3; Fig. 2). The 6p22.3 (rs34229995) and 7q36.1 (rs7781265) associations mark chromatin-regulating genes; rs34229995 is 2.2-kb telomeric to the 5′ of *JARID2* (jumonji, AT-rich interactive domain 2; Fig. 2) and rs7781265 localizing to intron 2 of *SMARCD3* (swi/swf-related, matrix-associated, actin-dependent regulator of chromatin, subfamily d, member 3; Fig. 2). The 10p12.1 (rs2790457) association localizes to intron 3 of the gene encoding WAC (WW domain-containing adaptor with coiled-coil region), which has recently been shown to be part of an extended autophagy network. The 20q13.13 (rs6066835) association mapped to intron 3 of *PREX1* (phosphatidylinositol-3, 4, 5-trisphosphate-dependent Rac exchange factor 1) (Fig. 2).
using all individuals in five of the six sample sets and observed no such relationships (Supplementary Tables 12 and 13). In addition, case-only analysis provided no evidence for associations between risk SNPs and cytogenetic MM subtype (Supplementary Table 14) or MM-specific overall survival (Supplementary Table 15). Collectively, these data are compatible with the risk variants having generic effects on MM development rather than tumour progression.

**Biological inference.** To the extent that they have been deciphered, many of the GWAS loci map to non-coding regions of the genome and influence gene regulation. In this respect, it is perhaps not surprising that none of the genes annotated by the GWAS signals we identify are somatically mutated in MM (Supplementary Fig. 2 and Supplementary Data 1). The direction of these eQTLs and meQTLs is compatible with the 10p12.1 1.12 signal encompassing an active promotor for PREX1 (Fig. 2).

DNA methylation plays a central role in epigenetic regulation of gene expression; however, meQTLs and cis-acting eQTLs do

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

| Location | SNP       | Position (bp) | Risk allele | RAF    | Data set | OR     | P-value |
|----------|-----------|---------------|-------------|--------|----------|--------|---------|
| 6p22.3   | rs34229995| 15,244,018    | G           | 0.029  | Discovery| 1.40   | 1.76 x 10^{-8} |
|          |           |               |             |        | Replication| 1.19   | 0.214  |
|          |           |               |             |        | Combined  | 1.37   | 1.31 x 10^{-8} |
|          |           |               |             |        | P_{het}  | 0.50   | 0%      |
| 6q21     | rs9372120 | 106,667,535   | G           | 0.218  | Discovery| 1.20   | 8.72 x 10^{-14} |
|          |           |               |             |        | Replication| 1.12   | 0.0147 |
|          |           |               |             |        | Combined  | 1.18   | 9.09 x 10^{-15} |
|          |           |               |             |        | P_{het}  | 0.93   | 0%      |
| 7q36.1   | rs7781265 | 150,950,940   | T           | 0.125  | Discovery| 1.20   | 1.82 x 10^{-7} |
|          |           |               |             |        | Replication| 1.15   | 0.0136 |
|          |           |               |             |        | Combined  | 1.19   | 9.71 x 10^{-9} |
|          |           |               |             |        | P_{het}  | 0.24   | 23%     |
| 8q24.21  | rs1948915 | 128,222,421   | C           | 0.345  | Discovery| 1.14   | 3.14 x 10^{-10} |
|          |           |               |             |        | Replication| 1.09   | 0.0283 |
|          |           |               |             |        | Combined  | 1.13   | 4.20 x 10^{-11} |
|          |           |               |             |        | P_{het}  | 0.34   | 11%     |
| 9p21.3   | rs2811710 | 21,991,923    | G           | 0.657  | Discovery| 1.14   | 6.50 x 10^{-10} |
|          |           |               |             |        | Replication| 1.18   | 4.02 x 10^{-5}  |
|          |           |               |             |        | Combined  | 1.15   | 1.72 x 10^{-13} |
|          |           |               |             |        | P_{het}  | 0.97   | 0%      |
| 10p12.1  | rs2790457 | 28,856,819    | G           | 0.739  | Discovery| 1.12   | 8.44 x 10^{-7} |
|          |           |               |             |        | Replication| 1.13   | 6.18 x 10^{-3}  |
|          |           |               |             |        | Combined  | 1.12   | 1.77 x 10^{-8} |
|          |           |               |             |        | P_{het}  | 0.94   | 0%      |
| 16q23.1  | rs7193541 | 74,664,743    | T           | 0.585  | Discovery| 1.12   | 1.14 x 10^{-8} |
|          |           |               |             |        | Replication| 1.17   | 4.79 x 10^{-4}  |
|          |           |               |             |        | Combined  | 1.13   | 5.00 x 10^{-12} |
|          |           |               |             |        | P_{het}  | 0.15   | 35%     |
| 20q13.13 | rs6066835 | 47,355,009    | C           | 0.083  | Discovery| 1.24   | 1.16 x 10^{-9} |
|          |           |               |             |        | Replication| 1.35   | 1.26 x 10^{-5}  |
|          |           |               |             |        | Combined  | 1.26   | 1.36 x 10^{-13} |
|          |           |               |             |        | P_{het}  | 0.072  | 43%     |

I^2, proportion of the total variation due to heterogeneity; OR, odds ratio; P_{het}, P-value for heterogeneity; RAF, risk allele frequency; SNP, single-nucleotide polymorphism. RAF is risk allele frequency across all cases and controls in the discovery set, where the risk allele is the allele corresponding to the estimated OR. Positions are based on NCBI build 37 of the human genome.
Figure 2 | Regional plots of association results and recombination rates for the newly identified risk loci for multiple myeloma. Results for 6p22.3 (rs34229995), 6q21 (rs9372120), 7q36.1 (rs7781265), 8q24.21 (rs1948915), 9p21.3 (rs2811710), 10p12.1 (rs2790457), 16q23.1 (rs7193541) and 20q13.13 (rs6066835). Plots (using visPig) show association results of both genotyped (triangles) and imputed (circles) SNPs in the GWAS samples and recombination rates. – Log10 P-values (y axes) of the SNPs are shown according to their chromosomal positions (x axes). The sentinel SNP in each combined analysis is shown as a large circle or triangle and is labelled by its rsID. The colour intensity of each symbol reflects the extent of LD with the top SNP, white ($r^2=0$) through to dark red ($r^2=1.0$). Genetic recombination rates, estimated using 1,000 Genomes Project samples, are shown with a light blue line. Physical positions are based on NCBI build 37 of the human genome. Also shown are the relative positions of genes and transcripts mapping to the region of association. Genes have been redrawn to show their relative positions; therefore, maps are not to physical scale. On the bottom is the chromatin-state segmentation track (ChromHMM) for lymphoblastoid cells using data from the HapMap ENCODE Project.
These sites are for known TFs such as nuclear factor-sensitivity common across multiple cell lines. The protein motifs at MM-related SNPs were observed in regions of DNase hypersensitivity for multiple cell lines. HaploReg showed that the majority of these TF binding or enhancer elements. We also assessed B-cell-specific chromatin dynamics using FANTOM5, which uses the pre-computed chromatin state data to identify patterns of local chromatin patterns, we analysed Hi-C data on the LCL cell line GM12878 (ref. 17), as a source of B-cell information (Supplementary Fig. 3). Loopping chromatin interactions and TADs were shown at 6q21 (rs9372120), 8q24.21 (rs1948915), 9p21.3 (rs2811710) and 20q13.13 (rs6066835), involving a number of genes with biological relevance to MM development. With the limitations of cell line data from LCL, which may not fully reflect MM biology, we demonstrated with MM RNA-sequencing data that gene expression within the 6q21 and 9p21.3 TADs were tightly correlated (P<2.0×10^-5), which is consistent with their co-regulation (Supplementary Table 17). Moreover, the region at 6q21 (rs9372120, MYG5) participates in PRC2 chromatin remodelling with the transcriptional repressor PRDM1 (Supplementary Fig. 3b). Similarly, the 8q24.21 region of association defined by rs1948915, which contains CCAT1 (colon cancer-associated transcript 1), interacts with MYC and distal upstream enhancer elements (Supplementary Fig. 3d).

To explore the epigenetic profile of association signals at each of the new MM risk loci, we used HaploReg and RegulomeDB to examine whether the sentinel SNPs and those in high LD (that is, r^2~0.8 in the 1,000 Genomes EUR reference panel) annotate putative transcription factor (TF) binding or enhancer elements. We also assessed B-cell-specific chromatin dynamics using FANTOM5, which uses the pre-computed chromatin state data for multiple cell lines. HaploReg showed that the majority of MM-related SNPs were observed in regions of DNase hypersensitivity common across multiple cell lines. The protein motifs at these sites are for known TFs such as nuclear factor-X, c-MYC, GATA, TCF4, POL24H8, CEBPB or POL2 (Supplementary Data 2). We examined for statistical evidence of enrichment in specific TF binding across the eight new and nine established risk loci using GM12878 data^18. Although of borderline significance and hypothesis generating, after correction for the 90 TFs assayed, there was evidence for enrichment of SP11 (alias PU.1), (P=0.0007, P_{adjusted}=0.063), which regulates PRDM1 and its downregulation is required for MM cell growth^19. Collectively, these observations are compatible with the identified risk SNPs mapping within regions of active chromatin state, which have a role in the B-cell cis-regulatory network.

**Discussion**

We have performed the largest GWAS of MM to date. We identified eight novel MM risk loci taking the total count to 17. Fully deciphering the functional impact of these SNP associations on MM development requires additional analyses. However, seven of the SNPs map intragenic to transcribed genes, which are relevant to MM or B-cell biology. Although a number of SNPs displayed an eQTL/meQTL in MM plasma cells, the absence of a relationship does not preclude the possibility of a subtle cumulative long-term relationship intrinsic to plasma cells or a predisposition through altered gene function in other cell types.

Studies in other cancers have shown that the multiple risk loci at 8q24.21 are enhancers interacting with MYC^20,21. As deregulation of MYC is a feature of MM, it is plausible that the susceptibility to MM has a similar mechanistic basis. Indeed, MYC promotes CCAT1 transcription by binding to its promoter, and in colorectal cancer the L isoform of CCAT1 has been shown to interact with the MYC promoter and distal upstream enhancer elements regulating MYC transcription^22. We have previously shown the MM risk SNP at 7p15.3 influences expression of CDCA7L, a binding partner of p75 potentiating MYC-mediated transformation. In addition to local interactions with CDKN2A/CDKN2B, the 9p21.3 region encompassing SNP rs2811710 interacts with the genomic region containing MTAP (methyithioadenosine phosphorylase). MTAP plays a major role in polyamine metabolism and deletion of MTAP is common in cancer, being closely linked to homozygous deletion of p16 (ref. 23).

**ATG5** at 6q21 is highly expressed in plasma cells and essential for autophagy and plasma cell survival^24. Strikingly, the same locus also contains the transcriptional repressor PRDM1 (formerly BLIMP1), which is key to the development of plasma cells from B cells and a determinant of plasma cell survival^25. The RFWD3 protein is an E3 ubiquitin ligase that positively regulates p53 stability by forming an RFWD3–MDM2–p53 complex, thereby protecting p53 from degradation by MDM2-mediated polyubiquitination^26. Variation at 16q23.1, defined with the SORL1 SNP rs4888262, has been shown to be associated with testicular cancer risk^27, suggesting a common genetic and biological basis to both associations.

**JARID2** functions as a transcriptional repressor through recruitment of Polycomb repressive complex 2 and has recently been identified as a regulator of haematopoietic stem cell function^28, and the 6p22.3-p21.31 region is commonly gained in MM tumours^29. Inhibition of JARID2 leads to loss of Polycomb binding and a reduction of histone H3 lysine-27 trimethylation levels on target genes. SMARCD3 recruits BAF chromatin remodelling complexes to specific enhancers. Although there is currently no evidence to implicate the transcriptional repressors JARID2 or SMARCD3 in terms of somatic mutation in MM, multiple genes including CDKN2A and TP53 are silenced by methylation in MM. Overexpression of histone methyltransferase and inactivating mutations in histone demethylase (UTX) typifies a subset of MM^30 and our findings add to the impact of chromatin remodelling genes on MM.

We have previously shown an association for MM at ULK4, a key regulator of mammalian target of rapamycin-mediated autophagy^4. We now suggest a more extensive set of associations involving ATG5 and WAC, and by virtue of the role of MYC in autophagy^31, CCAT1, CDCA7L, DNM3A and CBX7. Collectively, these data invoke deregulation of DNA methylation, telomere length, differentiation and autophagy, and immunoglobulin production as determinants of MM susceptibility.

Our findings provide further evidence for an inherited genetic susceptibility to MM. However, further studies are necessary to understand the biology behind these risk variants. We estimate that the currently identified risk SNPs for MM account for 20% of the heritable risk attributable to all common variation; hence, further GWAS-based studies in concert with functional analyses should lead to additional insights into MM biology. Importantly, such studies may inform the development of new therapeutic agents^32,33.

**Methods**

**Ethics.** Collection of patient samples and associated clinico-pathological information was undertaken with written informed consent and relevant ethical review board approval at respective study centres in accordance with the tenets of the Declaration of Helsinki, specifically for the Myeloma-IX trial by the Medical Research Council (MRC) Leukaemia Data Monitoring and Ethics committee (MREC 02/8/95, ISRCTN68441111), the Myeloma XI trial by the Oxfordshire Research Ethics Committee (MREC 17/09/09, ISRCTN49407852), HOVON65/GMMG-HD4 (ISRCTN 644552890; METC 13/01/2013), HOVON87/NMSG18.
Genome-wide association studies. The diagnosis of MM (ICD-10 C90.0) was established in accordance with World Health Organization guidelines. All samples from patients with MM for genotyping were obtained prior to treatment or at presentation. The meta-analysis was based on GWAS conducted in the Netherlands, the United Kingdom, Germany, Sweden/Norway, the United States and Iceland (Supplementary Tables 1 and 2).

The Dutch GWAS consisted of 608 cases (316 male). The cases were ascertained from three clinical trials: HOVON55/GMG-HD-ID, HOVON65/455289/GMG-HD-ID restricted to Dutch cases; n = 158), HOVON87/NMSGH (n = 292) and HOVON95/EMNO2 (n = 105) (ISRCTN64455289; HOVON87/NMSGH; HOVON87/NMSGH https://www.clinicaltrialsregister.eu/ctr-search/trial/2007/00047-34/BE and HOVON95/ EMNO2 https://www.clinicaltrialsregister.eu/ctr-search/trial/2009-017903-28(AT)).

DNA was extracted from venous blood samples and genotyped using Illumina Human OmniExpress-12 v1.0 arrays (Illumina, San Diego, USA). For controls, we used the B-PROOF data set (B-vitamins for the prevention of osteoporotic fractures). Controls were genotyped using Illumina OmniExpress Exome-v1-1 arrays.

The UK GWAS comprised 2,329 cases (1,060 male (post quality control (QC)); mean age at diagnosis: 64 years) recruited through the UK MRC Myeloma-IX and Myeloma-XI trials (ISRCTN68454111: Myeloma IX http://www.isrctn.com/search?q=ISRCTN68454111 and ISRCTN4907852: Myeloma XI http://www.isrctn.com/search?q=ISRCTN4907852). DNA was extracted from EDTA-venous blood samples (90% before chemotherapy) and genotyped using Illumina Human OmniExpress-12 v1.0 arrays (Illumina). For controls, we used publicly accessible data generated by the Wellcome Trust Case Control Consortium from www.isrctn.com/search?q=ISRCTN49407852). DNA was prepared from EDTA-venous blood samples (90% before chemotherapy) and genotyped using Illumina Human OmniExpress-12 v1.0 arrays (Illumina). For controls, we used genotype data on 2,107 healthy individuals, enroled into the Heinz Nixdorf Recall (HNR) study genotyped using either Illumina HumanOmni-Quad_v1 or 1428 OmniExpress-12 v1.0 arrays.

The Swedish/Norwegian GWAS was based on 1,668 and 157 MM patients from the Swedish/Norwegian Myeloma Biobank for Myeloma (Trondheim, Norway) and the Norwegian Biobank for Myeloma (Trondheim, Norway), respectively. Genotyping was performed using Illumina Human OmniExpress-12 v1.0 arrays (Illumina). For controls, we used genotype data on 2,107 healthy individuals, enroled into the Heinz Nixdorf Recall (HNR) study genotyped using either Illumina HumanOmni-Quad_v1 or 1428 OmniExpress-12 v1.0 arrays.

The Danish GWAS was based on 1,076 newly diagnosed patients treated at the UAMS Myeloma Institute for Research and Therapy (NCT00083551: Total therapy II https://clinicaltrials.gov/ct2/show/NCT00083551; NCT00081939: Total therapy III https://clinicaltrials.gov/ct2/show/NCT00081939; NCT0572619: Total therapy 3B https://clinicaltrials.gov/ct2/show/NCT00572619; and NCT0734877: Total therapy 3).

The Icelandic GWAS comprised 480 MM cases identified from the nationwide Icelandic Cancer Registry. Samples were genotyped using Illumina microarrays.

Analysis of GWAS. The Swedish/Norwegian GWAS has been previously published in its entirety with a full description of QC. Adopting the same standard, quality-control measures were applied to the UK, German, US and the Netherlands GWAS. Specifically, we excluded individuals with low call rate (<95%) and those found to have non-European ancestry on the basis of HapMap version 2 CEU, JPT/CHB and YRI population reference data (Supplementary Fig. 4). For first-degree relative pairs, we excluded the control or the individual with the lower call rate. SNPs with a call rate <95% were excluded as were those with a MAF <0.01 or displaying significant deviation from Hardy–Weinberg equilibrium (that is, P < 10−5). Post QC, the 5 GWAS provided genotype data on 6,839 cases and 22,221 controls. GWAS data were imputed for all scans for >10 million SNPs using 1,000 Genomes Project (phase 1 integrated release 3, March 2012) and UK10k data (ALSAPAC, EGAS0000100009/EGAD0000100195 and TwinsUK EGAS0000100009/EGAD0000100194 studies only) as reference in conjunction with UK10K and IMPUTE2 v2.3 software (University of Arkansas for Medical Sciences Institutional Review Board (IRB 20077), Lund University Ethical Review Board (2013/54) and Icelandic Data Protection Authority (2,001,010,157 and National Bioethics Committee 01/015).

Meta-analysis. We performed association testing in the discovery sets separately and then combined the results for the 12.4 million variants. We assessed the fidelity of imputation through the concordance between imputed and directly genotyped SNPs. The imputation performance of GWAS samples (Supplementary Tables 6 and 7). Meta-analysis was undertaken using the inverse-variance approach under a fixed-effects model implemented in META v1.6 (ref. 39). Cochran's Q-statistic was calculated, to test for heterogeneity, and the I2 statistic measured, to quantify the proportion of the total variation due to heterogeneity. Meta-analysis summary statistics and LD controls for a reference panel of 1,000 Genomes Project combined with UK10K, we used GCTA to perform conditional association analysis. Association statistics were calculated for all SNPs conditioning on the top SNP in each loci showing genome-wide significance. This is performed in a step-wise manner.

Replication genotyping. To validate promising associations, we analysed four case–control series from the United Kingdom, Germany, Denmark and Sweden/Norway.

The UK replication comprised 812 MM cases (412 male) ascertained through the UK 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 the National Lung Cancer Pre-disposition Study (n = 536) (ref. 42) and National Study of Colorectal Cancer Genetics (n = 574) (ref. 43). All cases and controls were UK residents.

The German replication series comprised 1,149 cases collected by the German Myeloma Study Group (Deutsche Studiengruppe Myelom, DSGM), GMMC University Clinic, Heidelberg, and University Clinic, Ulm (676 male, mean age at diagnosis 57.6 years, s.d. 9.8). Controls comprised of 1,582 healthy German blood donors recruited between 2004 and 2007 by the Institute of Transfusion Medicine and Immunology, University of Mannheim, Germany (885 male, mean age 55.8 years, s.d. 10.0).

The Swedish/Norway and Danish replication series comprised 223 MM cases from the Swedish National Myeloma Biobank and 363 MM cases from the University Hospital of Copenhagen. As controls for these respective replication sets, we analysed 1,285 Swedish blood donors and 826 individuals from Denmark and Skåne County, Sweden (the southernmost part of Sweden adjacent to Denmark).

Replication genotyping was performed using allele-specific PCR KASPar chemistry (LGC, Hertfordshire, UK; UK replication series). Primers, probes and conditions used are available on request. Call rates for SNP genotypes were >95% in each of the replication series. The quality of genotyping in all assays was assessed by measuring 1–10% duplicates (showing a concordance of >99%) and at least two negative controls for each centre. Technical artefacts were excluded by cross-platform validation of 96 samples and sequencing of a set of 96 randomly selected samples from each case and control series confirmed genotyping accuracy. Concordance of >99% demonstrated robust performance.

Translocation detection and mutation analysis. Carotyopon was used for cytogenetic studies of MM cells and standard criteria for the definition of a clone were applied. The frequency of somatic mutation in genes annotated by GWAS signals was derived from tumour whole-exome sequencing of 463 Myeloma XI trial patients.
Association between genotype and patient outcome. To examine the relationship between SNP genotype and patient outcome, we analysed GWAS data on four of the patient cohorts3–5, specifically (i) 1,165 cases from the UK MRC Myeloma IX trial (UK-GWAS); (ii) 877 MM cases from the UK MRC Myeloma-XI trial (UK-GWAS); (iii) 511 of the patients recruited to the German GWAS; and (iv) 703 MM cases in the UAMS Myeloma Institute for Research and Therapy GWAS (USA GWAS). Clinical trial information on these patients has been previously reported41–43. The analysis was performed following the guidelines described44–46. Our Cox regression analysis was used to derive genotype-specific hazard ratio and associated 95% confidence intervals. Meta-analysis was performed under a fixed-effects model (Supplementary Table 15).

eQTL analysis. We performed an eQTL analyses using Affymetrix Human Genome U133 2.0 Plus Array data for plasma cells from 183 MRC Myeloma IX trial patients29, 658 Heidelberg patients and 608 US patients as recently described. Briefly, GER, UK and US data were separately pre-processed and analysed using a Bayesian approach to probabilistic estimation of expression residuals to infer broad variance components, thus accounting for hidden determinants influencing global expression such as copy number, translocation status and batch effects47–49. The association between genotype of the sentinel variant and gene expression of genes within 500 kbp either side was evaluated based on the significance of linear regression coefficients. We pooled data from each study under a fixed-effects model controlling for false discovery rate (FDR) calling significant associations with a FDR < 0.05. In addition, we queried publicly available eQTL messenger RNA expression data using MuTHER and the Blood eQTL browser. MuTHER contains expression data on lymph skin and adipose tissue with a European GWAS41,66,67. The Blood eQTL browser contains expression data from 5,311 non-transformed peripheral blood samples54.

meQTL analysis. We performed cis-meQTL analysis using Illumina 450K methylation array data on plasma cells from 384 MRC Myeloma XI trial patients. As with analysis of MM expression (eQTL) data, we inferred hidden determinants influencing global methylation. The genetic association was tested under an additive model between each SNP and each normalized methylation probe, adjusting for plate of the methylation-based principal component analysis score. Controlling for a FDR of 0.05 across the 338,456 methylation traits required a P-value for association to be < 4.0 x 10⁻⁵.

ENCODE and chromatin state dynamics. Risk SNPs and their proxies (that is, rs² > 0.8 in the 1,000 Genomes EUR reference panel) were annotated for putative functional effect using HaploReg v3 (ref. 55), RegulomeDB56 and SeattleSeq57 annotation. These servers make use of data from ENCODE58, genomic evolutionary rate profiling59 conservation metrics, combined annotation dependent evolutionary scores60 and PolyPhen score61. We obtained an overlap of associated SNPs with predicted enhancers using the FANTOM5 enhancer atlas62 and searched for overlap with ‘super-enhancer’ regions using data from Hnisz et al.63, restricting our analysis to GM12878.

To formally examine for enrichment in specific TF binding across risk loci, we adopted the method described by Dalton et al.64. Briefly, for each risk locus we derived a credible set of SNPs with a 99% probability of containing the causal SNP; posterior probability for each SNP being computed from its Bayes factors. SNPs were ranked by their posterior probability and included so that the cumulative posterior probability for association was >0.99. Binding sites for 90 TF in GM12878 were obtained from ENCODE. For each TF the total posterior probability over all credible set SNPs overlapping all binding sites was calculated. A null distribution was generated by randomly relocating each binding site up to 100 kb from its original location. For these perturbed sites, the total posterior probability over all overlapping SNPs was calculated. This process was repeated 10,000 times and enrichment P-values calculated as the fraction of permutations where the total cumulative posterior probability was greater than for the unperturbed binding sites.

Hi-C data and definition of topological domains at risk loci. Hi-C data was used to map the candidate causal SNPs to chromosomal TADs and identify patterns of relevant, local chromatin interactions. We made use of publicly available raw Hi-C data on GM12878 cells17. Valid Hi-C pairs were identified aligning raw reads to the reference genome using Burrows-Wheeler alignment (BWA), matching pairs of reads and filtering for biases. Bonfide Hi-C ditags were allocated to a contact matrix, with a predefined, uniform resolution of 5 kb. We corrected for experimental bias using the matrix balancing approach65. We inferred TADs from the contact matrix by means of the arrowhead algorithm for domain detection as previously proposed66.

To investigate whether genes within TADs are co-regulated, we obtained RNAseq transcript counts from 66 MM cell lines from the Keat’s lab Data Repository (http://www.keatlab.org/data-repository)67. We performed pairwise correlation by calculating the Pearson’s product–moment correlation coefficient of the transcript counts for all pairs of genes within respective TADs.

Heritability analysis. We used Genome-wide Complex Trait Analysis to estimate the polygenic variance ascribable to all genotyped and imputed GWAS SNPs simultaneously for the UK and German GWAS41,66,67. SNPs were excluded based on low MAF, poor imputation and poor HWE. Principal components were included as covariates in the heritability analysis of the German data. As previously advocated when calculating the heritability of a disease such as cancer we used the lifetime risk68, which for MM is estimated to be 0.007 for the UK population (http://www.cancerresearchuk.org/cancer-info/cancerstats/types/myeloma/myeloma-incidence/uk-multiple-myeloma-incidence-statistics#Lifetime) and 0.006 for the German population. We estimated the heritability explained by risk SNPs identified by GWAS as located within regions associated with MM. Meta-analysis of heritability estimates from UK and German GWAS data sets was performed under a standard fixed-effects model.

Data availability. SNP genotyping data that support the findings of this study have been deposited in Gene Expression Omnibus with accession codes GSE21349, GSE19784, GSE24080, GSE2658 and GSE15695; in the European Genome-phenome Archive (EGA) with accession code EGAS000000000001; in the European Bioinformatics Institute (Part of the European Molecular Biology Laboratory) (EMBL-EBI) with accession code E-MTAB-362 and E-TABM-1138; and in the database of Genotypes and Phenotypes (dbGaP) with accession code phs000107v07.

Expression data that support the findings of this study have been deposited in GEO with accession codes GSE21349, GSE2658, GSE1161 and EMBL-EBI with accession code E-MTAB-2299.

Whole-exome sequence data that support the findings of this study have been deposited in EGA with accession code EGAS00000101147.

Transcription profiling data from MuTHER studies that support the findings of this study have been deposited in EMBL-EBI with accession code E-TABM-1104. Data from Blood eQTL have been deposited in EMBL-EBI with accession codes E-TABM-1036, E-MTAB-945 and E-MTAB-1708.

The remaining data are contained within the paper and Supplementary Files or available from the author upon request.

References
1. Moor, G. J. et al. Inherited genetic susceptibility to multiple myeloma. Leukemia 28, 518–524 (2014).
2. Chubb, D. et al. Common variation at 3q26.2, 6p21.33, 17p11.2 and 22q13.1 influences multiple myeloma risk. Nat. Genet. 45, 1221–1225 (2013).
3. Weinhold, N. et al. The CCND1 c.870G>A polymorphism is a risk factor for t(11;14)(q13;xq32) multiple myeloma. Nat. Genet. 45, 522–525 (2013).
4. Broderick, P. et al. Common variation at 3p22.1 and 7p15.3 influences multiple myeloma risk. Nat. Genet. 44, 58–61 (2012).
5. Swaminathan, B. et al. Variants in EZH2 influencing immunoglobulin levels associate with multiple myeloma. Nat. Commun. 6, 7213 (2015).
6. Mitchell, J. S. et al. Implementation of genome-wide complex trait analysis to quantify the heritability in multiple myeloma. Sci. Rep. 5, 12473 (2015).
7. Erickson, S. W. et al. Genome-wide scan identifies variant in 2q12.3 associated with risk for multiple myeloma. Blood 124, 2001–2003 (2014).
8. Genomes Project, C. et al. A map of human genome variation from population-scale sequencing. Nature 467, 1061–1073 (2010).
9. Huang, J. et al. Improved imputation of low-frequency and rare variants using the UK10K haplotype reference panel. Nat. Commun. 6, 8111 (2015).
10. Guðbjartsson, D. F. et al. Large-scale whole-genome sequencing of the Icelandic population. Nat. Genet. 47, 435–444 (2015).
11. Fletcher, O. & Houlston, R. S. Architecture of inherited susceptibility to common cancer. Nat. Rev. Cancer 10, 353–361 (2010).
12. Gerhan, J. J. et al. Genome-wide association study identifies multiple susceptibility loci for diffuse large B cell lymphoma. Nat. Genet. 46, 1233–1238 (2014).
13. 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. 42, 1126–1130 (2010).
14. Crowther-Swanepoel, D. et al. Common variants at 2q37.3, 8q24.21, 15q23.1 and 16q24.1 influence chronic lymphocytic leukemia risk. Nat. Genet. 42, 132–136 (2010).
15. Sherborne, A. L. et al. Variation in CCR2A2 at 9p21.3 influences childhood acute lymphoblastic leukemia risk. Nat. Genet. 42, 492–494 (2010).
16. Joachim, J., Wirth, M., McKnight, N. C. & Tooze, S. A. Coiling up with SCOC and Bona fide retroviral integration to examine the chromatin organisation of common cancer. Nat. Rev. Cancer 14, 1626–1630 (2014).
17. Rao, S. S., et al. A 3D map of the human genome at kilobase resolution reveals organization of the core transcriptional machinery. Cell 159, 1665–1680 (2014).
18. Gaulton, K. J. et al. Genetic fine mapping and genomic annotation defines causal mechanisms at type 2 diabetes susceptibility loci. Nat. Genet. 47, 1415–1425 (2015).
41. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Mol. Immunol.* 52, 5237–5248 (2015).

42. Nelson, M. R. et al. The support of human genetic evidence for approved drug indications. *Nat. Genet.* 47, 856–860 (2015).

43. Ocio, E. M., Mateos, M. V., Maiso, P., Pandiella, A. & San-Miguel, J. F. New PRMT5 arginine methyltransferase in cancer cells. *Nature* 526, 112–117 (2015).

44. Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genome imputation method for the next generation of genome-wide association studies. *PLoS Genet.* 5, e1000529 (2009).

45. 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.* 39, 906–913 (2007).

46. Clayton, D. G. et al. Population structure, differential bias and genomic control in a large-scale, case-control association study. *Nat. Genet.* 37, 1243–1247 (2005).

47. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.* 2, e190 (2006).

48. Liu, J. Z. et al. Meta-analysis and imputation refines the association of 15q25 SMARCA4 with bladder cancer risk. *Hum. Mol. Genet.* 22, 5237–5248 (2013).

49. Morgan, G. J. et al. Cyclophosphamide, thalidomide, and dexamethasone as induction therapy for newly diagnosed multiple myeloma patients destined for autologous stem-cell transplantation: MRC Myeloma IX randomized trial results. *Haematologica* 97, 442–450 (2012).

50. Morgan, G. J. et al. Long-term follow-up of MRC Myeloma IX trial: survival outcomes with biphosphonate and thalidomide treatment. *Clin. Cancer Res.* 19, 6030–6038 (2013).

51. Johnson, D. C. et al. Genome-wide association study identifies variation at 6q25.1 associated with survival in multiple myeloma. *Nat. Commun.* 7, 10290 (2016).

52. Stegle, O., Parts, L., Piipari, M., Winn, I. & Durbin, R. Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. *Nat. Protoc.* 7, 500–507 (2012).

53. Grundberg, E. et al. Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nat. Genet.* 47, 1001–1003 (2015).

54. Westra, H. J. et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat. Genet.* 45, 1238–1243 (2013).

55. Ward, L. D. & Kellis, M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* 40, D930–D934 (2012).

56. Chapman, M. A. et al. Initial genome sequencing and analysis of multiple myeloma. *Nature* 471, 467–472 (2011).

57. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. Genome-wide complex trait analysis (GCTA): methods, data analyses, and interpretations. *Methods Mol. Biol.* 1019, 215–236 (2013).

58. Yang, J. et al. Genome partitioning of genetic variation for complex traits using common SNPs. *Nat. Genet.* 43, 519–525 (2011).

59. Hnisz, D. et al. Super-enhancers in the control of cell identity and disease. *Cell* 155, 934–947 (2013).

60. Rager, J. et al. Capture Hi-C identifies the chromatin interactome of colorectal cancer risk loci. *Nat. Commun.* 6, 6178 (2015).

61. Chapman, M. A. et al. Initial genome sequencing and analysis of multiple myeloma. *Nature* 471, 467–472 (2011).

62. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. Genome-wide complex trait analysis (GCTA): methods, data analyses, and interpretations. *Methods Mol. Biol.* 1019, 215–236 (2013).

63. Yang, J. et al. Genome partitioning of genetic variation for complex traits using common SNPs. *Nat. Genet.* 43, 519–525 (2011).

64. Lu, Y. et al. Most common ‘sporadic’ cancers have a significant germline genetic component. *Hum. Mutat.* 33, 6112–6118 (2014).

65. Lee, S. H. et al. Estimation and partitioning of polygenic variation captured by common SNPs for Alzheimer’s disease, multiple sclerosis and endometriosis. *Hum. Mutat.* 32, 832–841 (2011).

66. Scales, M., Rager, J., Mihalek, G., Houlston, R. S. & Henzri, M. Y. VisPI-LIE--a web tool for producing multi-region, multi-task, multi-scale plots of genetic data. *PLoS ONE* 9, e107497 (2014).

Acknowledgements

In the United Kingdom, Myeloma UK and Bloodwise provided principal funding. Additional funding was provided by Cancer Research UK (C1298/AA362 supported by the Bobby Moore Fund) and The Rosetrees Trust. This study made use of genotyping data on the 1958 Birth Cohort generated by the Wellcome Trust Sanger Institute (http://www.wtccc.org.uk). We are grateful to all investigators who contributed to NSCCG and GELCAPS, from which controls in the replication were drawn. We also thank the staff of the CRU University of Leeds and the NCRI haematology Clinical Studies Group. The US GWAS was supported by a grant from the National Institutes of Health (P01CA055819). The German study was supported by the Dietmar-Hopp-Stiftung, Germany, the German Cancer Aid (110,131), the German Ministry of Education and Science (CLIMMICS 01ZX1309), the German Research Council (DFG Project SI 236/8–1, SI236/9–1, 1ER 155/6–1 and the DFG CRU 216) and the Multiple Myeloma Research Foundation. The patients were collected by the GMMG and DSSM studies. The German GWAS was made up of genotyping data from the population-based HNR study, which is supported by the Heinz Nixdorf Foundation (Germany). The genotyping of the Illumina HumanOmni1-Quad ReadChips was performed by the members of the HNR subjects was financed by the German Center for Neurodegenerative Disorders (DZNE), Bonn. We are grateful to all investigators who contributed to the generation of this data set. The German replication controls were collected by Peter Bugert, Institute of Transfusion Medicine and Immunology, Medical Faculty Mannheim, Heidelberg University; German Red Cross Blood Services Baden-Württemberg/Germany. This work was supported by research grants from the Swedish Foundation for Strategic Research (KF10-0009), the Marianne and Marcus Wallenberg Foundation (2010.0112), the Knut and Alice Wallenberg Foundation (2012.0193), the Swedish Research Council (2012–1753), the Royal Swedish Academy of Sciences, and the U.S. National Institutes of Health (AR055641-02A1). The Natural Environment Research Council (NERC) and the Stroke Association (SA) supported the research on the epidemiology of multiple myeloma.
Swedish Academy of Science, ALF grants to the University and Regional Laboratories (Labmedicin Skåne), the Sv-Inger and Per-Erik Andersson Foundation, the Medical Faculty at Lund University and the Swedish Society of Medicine. We thank Jörgen Adolfsson, Tomas Axelsson, Anna Collin, Ildikó Frigyesi, Patrik Magnusson, Bertil Johansson, Jan Westin and Helga Ogmundsdóttir for their assistance. We are indebted to the clinicians who contributed samples to Swedish, Icelandic, Norwegian and Danish biobanks. We are indebted to the patients and other individuals who participated in the project.

Author contributions
R.S.H. designed the study. R.S.H. drafted the manuscript with contributions from K.H., G.J.M., B.N., N.W., J.S.M. and M.D. J.S.M. performed principal statistical and bioinformatic analyses. N.L., D.C.J., M.H., G.M. and O.L. performed additional bioinformatics analyses. P.B. coordinated UK laboratory analyses. N.L. performed genotyping and sequencing of UK samples. D.C.J. managed and prepared Myeloma IX and Myeloma XI Case Study DNA samples. M.K., G.J.M., F.E.D., W.A.G. and G.H.J. performed ascertainment and collection of Case Study samples. B.A.W. performed principal statistical and bioinformatic analyses. F.M.R. performed UK fluorescence in situ hybridization analyses. H.G., U.B., J.H., J.N., and N.W. coordinated and managed Heidelberg samples. C.L. and H.E. coordinated and managed Ulm/Würzburg samples. M.A. performed German genotyping. P.H. and M.M.N. performed GWAS of German cases and controls. B.C. and M.I.d.S.F. carried out statistical analysis. K.H. coordinated the German part of the project. M.M.N. generated genotype data from the Heinz-Nixdorf recall study. M.H. and B.N. coordinated the Swedish/Norwegian part of the project. M.A. and B.-M.H. performed data analysis. B.S., M.I., E.I., S.I., C.H., A.-K.W., U.-H.M., H.N., S.N., A.V., U.V., A.W., I.T. and U.G. performed sample acquisition, sample preparation, clinical data acquisition and additional data analyses of Sweden/Norway samples. In Iceland, G.T. and D.F.G performed statistical analysis. S.Y.K. provided clinical information. T.R. performed additional analyses. U.T. and K.S. performed project oversight. M.K., P.S., A.B. and R.K. performed ascertainment and collection of Case Study samples.

Additional information
Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: G.T., P.S., G.M., D.F.G., T.R., K.S. and U.T. are employed by deCode Genetics/Amgen Inc. The remaining authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Mitchell, J. S. et al. Genome-wide association study identifies multiple susceptibility loci for multiple myeloma. Nat. Commun. 7:12050 doi: 10.1038/ncomms12050 (2016).