

ARTICLE

Received 18 Sep 2014 | Accepted 14 May 2015 | Published 22 Jul 2015

DOI: 10.1038/ncomms8539

# Genome-wide association study identifies variants at 16p13 associated with survival in multiple myeloma patients

Elad Ziv<sup>1,\*</sup>, Eric Dean<sup>2,\*</sup>, Donglei Hu<sup>1</sup>, Alessandro Martino<sup>3</sup>, Daniel Serie<sup>4</sup>, Karen Curtin<sup>5</sup>, Daniele Campa<sup>3</sup>, Blake Aftab<sup>6</sup>, Paige Bracci<sup>7</sup>, Gabriele Buda<sup>8</sup>, Yi Zhao<sup>5</sup>, Jennifer Caswell-Jin<sup>1</sup>, Robert Diasio<sup>9</sup>, Charles Dumontet<sup>10</sup>, Marek Dudziński<sup>11</sup>, Laura Fejerman<sup>1</sup>, Alexandra Greenberg<sup>12,13</sup>, Scott Huntsman<sup>1</sup>, Krzysztof Jamrozik<sup>14</sup>, Artur Jurczyszyn<sup>15</sup>, Shaji Kumar<sup>16</sup>, Djordje Atanackovic<sup>5</sup>, Martha Glenn<sup>5</sup>, Lisa A. Cannon-Albright<sup>5</sup>, Brandt Jones<sup>5</sup>, Adam Lee<sup>9</sup>, Herlander Marques<sup>17</sup>, Thomas Martin<sup>6</sup>, Joaquin Martinez-Lopez<sup>18</sup>, Vincent Rajkumar<sup>16</sup>, Juan Sainz<sup>19,20</sup>, Annette Juul Vangsted<sup>21</sup>, Marzena Wątek<sup>22</sup>, Jeffrey Wolf<sup>6</sup>, Susan Slager<sup>23</sup>, Nicola J. Camp<sup>5,†</sup>, Federico Canzian<sup>3,†</sup> & Celine Vachon<sup>13,†</sup>

Here we perform the first genome-wide association study (GWAS) of multiple myeloma (MM) survival. In a meta-analysis of 306 MM patients treated at UCSF and 239 patients treated at the Mayo clinic, we find a significant association between SNPs near the gene *FOPNL* on chromosome 16p13 and survival (rs72773978;  $P = 6 \times 10^{-10}$ ). Patients with the minor allele are at increased risk for mortality (HR: 2.65; 95% CI: 1.94–3.58) relative to patients homozygous for the major allele. We replicate the association in the IMMENSE cohort including 772 patients, and a University of Utah cohort including 318 patients (rs72773978  $P = 0.044$ ). Using publicly available data, we find that the minor allele was associated with increased expression of *FOPNL* and increased expression of *FOPNL* was associated with higher expression of centrosomal genes and with shorter survival. Polymorphisms at the *FOPNL* locus are associated with survival among MM patients.

<sup>1</sup>Division of General Internal Medicine, Department of Medicine, Institute for Human Genetics, Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, California 94143, USA. <sup>2</sup>Sutter Medical Center of Santa Rosa, Santa Rosa, California 95403, USA. <sup>3</sup>Genomic Epidemiology Group, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany. <sup>4</sup>Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic, Jacksonville, Florida 32224, USA. <sup>5</sup>Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, Utah 84108, USA. <sup>6</sup>Division of Hematology, Department of Medicine, Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, California 94143, USA. <sup>7</sup>Department of Epidemiology and Biostatistics, Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, California 94143, USA. <sup>8</sup>Department of Oncology, Transplants and Advanced Technologies, Section of Hematology, Pisa University Hospital, Pisa, Italy. <sup>9</sup>Department of Molecular Pharmacology and Experimental Therapeutics, College of Medicine, Mayo Clinic, Rochester, Minnesota 55902, USA. <sup>10</sup>INSERM UMR 1052/CNRS 5286, Laboratoire de Cytologie Analytique, Faculté de Médecine Rockefeller, Université Claude Bernard Lyon I, Lyon 69008, France. <sup>11</sup>Department of Hematology, Rzeszow Regional Hospital, Rzeszow 35-301, Poland. <sup>12</sup>Center for Translational Science Activities, Mayo Clinic, Rochester, Minnesota 55905, USA. <sup>13</sup>Division of Epidemiology, Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota 55905, USA. <sup>14</sup>Department of Hematology, Medical University of Lodz, Lodz 93-510, Poland. <sup>15</sup>Department of Hematology, Cracow University Hospital, Cracow 31-501, Poland. <sup>16</sup>Division of Hematology, Department of Internal Medicine, Mayo Clinic, Rochester, Minnesota 55905, USA. <sup>17</sup>Life and Health Sciences Research Institute (ICVS), School of Health Sciences, University of Minho, Braga 4710-243, Portugal. <sup>18</sup>Hematology Service, CRIS facility for Hematological research, Hospital Universitario 12 de Octubre, Universidad Complutense, Madrid 28041, Spain. <sup>19</sup>Genomic Oncology Area, GENYO, Centre for Genomics and Oncological Research: Pfizer/University of Granada/Andalusian Regional Government, Granada 18014, Spain. <sup>20</sup>Department of Hematology, Virgen de las Nieves University Hospital, Granada 18014, Spain. <sup>21</sup>Department of Hematology, Rigshospitalet and Roskilde Hospital, Copenhagen University, 2100 Copenhagen, Denmark. <sup>22</sup>Department of Hematology, Holycross Cancer Center, Kielce 24-734, Poland. <sup>23</sup>Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic College of Medicine, Rochester, Minnesota 55905, USA. \* These authors contributed equally to this work. † These authors jointly supervised this work. Correspondence and requests for materials should be addressed to E.Z. (email: elad.ziv@ucsf.edu).

**M**ultiple myeloma (MM) is an incurable haematological malignancy of plasma cells. Approximately 22,000 new cases are diagnosed each year in the United States and over 10,000 deaths occur annually<sup>1</sup>. Family history is a strong risk factor for MM<sup>2</sup>. Recent genome-wide association studies (GWAS) reported eight loci associated with susceptibility to MM<sup>3–5</sup>.

A variety of clinical features and biomarkers are associated with MM prognosis<sup>6</sup>. Chromosomal abnormalities are also associated with prognosis; deletions at 17p, 13q, amplifications at 1q and translocations t(4;14) and t(14;16) have been associated with a poor prognosis, while hyperdiploidy is associated with a favourable prognosis<sup>7–9</sup>. Gene expression signatures of the myeloma cells also predict survival<sup>10–14</sup>.

Germline genetic variants are associated with survival among patients with oesophageal<sup>15</sup>, breast<sup>16,17</sup>, pancreatic<sup>18,19</sup> and small cell lung cancer<sup>20</sup>. We performed the first GWAS of MM survival, by conducting a meta-analysis of two studies from the University of California San Francisco (UCSF) and the Mayo Clinic. We found a locus on chromosome 16 associated with survival. We replicated the findings in the International Multiple Myeloma rESEarch (IMMEnSE) consortium and the University of Utah cohort. The top single-nucleotide polymorphisms (SNPs) were at the *FOPNL* locus. Using publicly available data of gene expression from peripheral blood of normal individuals, we found that the risk alleles at the top SNPs were associated with increased expression of *FOPNL*. We also found that increased expression of *FOPNL* was associated with higher 'centrosome index', a gene expression correlate of centrosome amplification in MM cells, which has been associated with poor survival.

## Results

**Identifying a locus for survival.** We performed a GWAS of overall survival among MM patients in cohorts (Table 1) from UCSF ( $n = 306$ ) and Mayo Clinic ( $n = 239$ ) separately. One locus mapping to chromosome 16p13.11 (Hg19) showed a suggestive association in both the UCSF ( $P = 8.4 \times 10^{-7}$ ; proportional hazards model) and the Mayo Clinic studies ( $P = 1.1 \times 10^{-4}$ ; proportional hazards model). In a meta-analysis of these GWAS,

the locus was genome-wide significant (Fig. 1a) with the strongest evidence at two SNPs in perfect linkage disequilibrium rs72773978 and rs117863986 ( $P = 6.0 \times 10^{-10}$  for both; meta-analysis  $P$  value is calculated using inverse variance-based weighting). We found no significant deviation from the proportional hazards assumption for the top SNP in either the UCSF ( $P = 0.74$ ;  $P$  values calculated by testing whether scaled Schoenfeld residuals vary with time) or Mayo clinic studies ( $P = 0.95$ ). We identified 131 SNPs at this locus associated with survival at  $P < 5 \times 10^{-8}$  (Supplementary Data 1); these SNPs had ~5–7% minor allele frequency (MAF) and were in tight linkage disequilibrium ( $r^2 > 0.8$ ) with the top SNPs (Fig. 1b). Of the 131 top SNPs, 17 were genotyped in the UCSF data set and 1 was genotyped in the Mayo Clinic data set (Supplementary Data 1). The remaining SNPs were imputed but had very high imputation quality scores (Information  $> 0.9$  or  $r^2 > 0.9$ ). We directly genotyped eight additional SNPs in the Mayo clinic data set, including one of the top two SNPs, rs117863986, and found consistently strong levels of association with the genotyped SNPs (rs117863986 hazard ratio (HR): 2.26; 95% confidence interval (CI): 1.46–3.40;  $P = 0.00021$ ; proportional hazards model) and other SNPs (Supplementary Table 1).

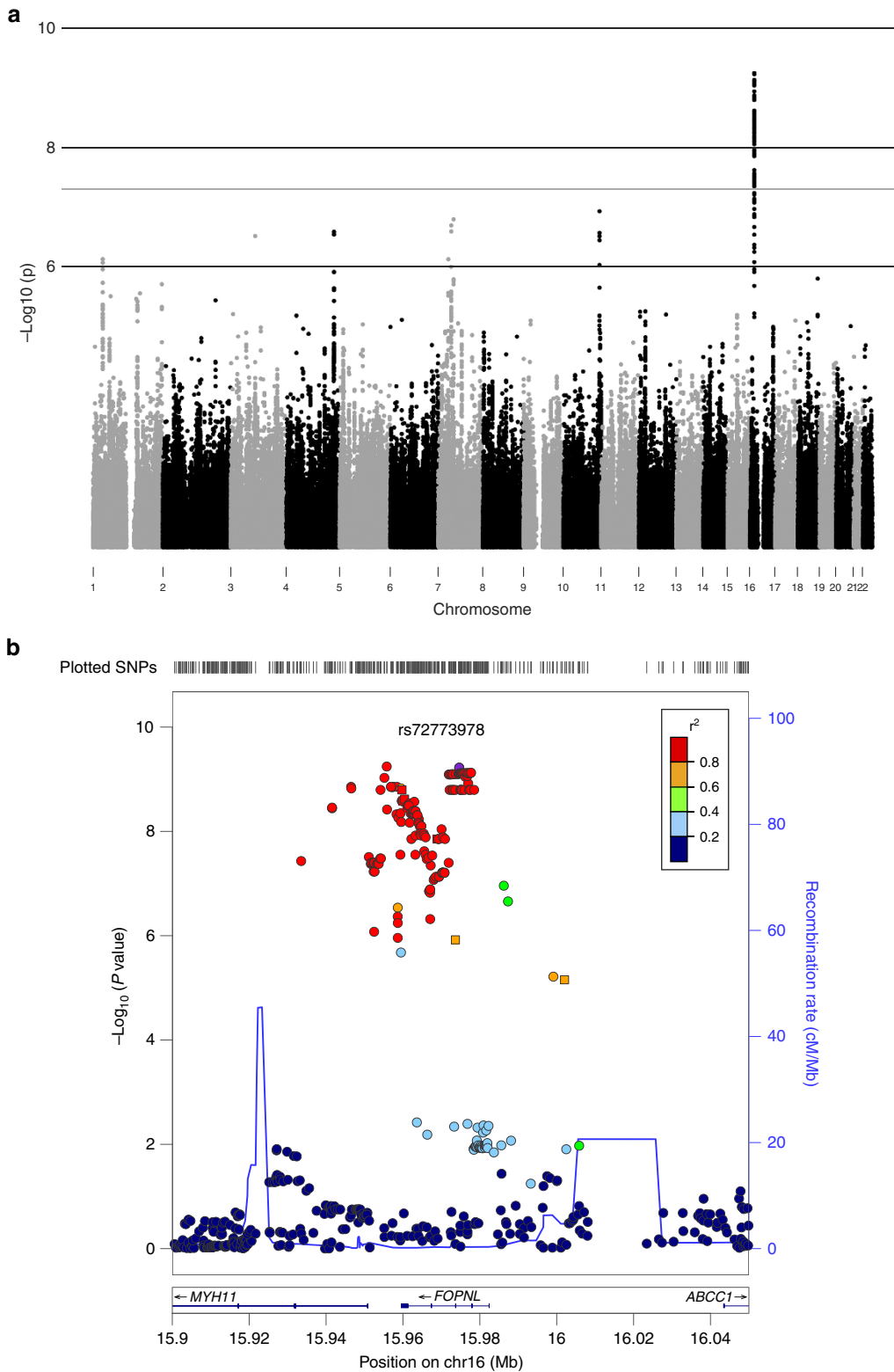
Analysis of the genome-wide distribution of association statistics (Supplementary Fig. 1a) revealed minimal deviation from the expectation under the null ( $\lambda = 1.002$ ). After removing SNPs from a 200-kB region around the top locus on chromosome 16, we found no evidence for additional signal genome wide (Supplementary Fig. 1b), although some other loci had some suggestive signals with  $P$  values  $5 \times 10^{-7}$ – $1 \times 10^{-7}$  (Supplementary Data 1).

The UCSF study had a median time of 7.6 months (interquartile range 5.7–8.9 months) between the date of diagnosis and the date of ascertainment. Therefore, we considered whether this delay affected our results. First, we adjusted for the time difference between the date of diagnosis and ascertainment in the proportional hazards models and found no attenuation (Supplementary Table 2). We also considered models stratified by the delay between date of diagnosis and ascertainment

**Table 1 | Clinical characteristics of patients in discovery and replication studies.**

	UCSF (N = 306)	Mayo Clinic (N = 239)	IMMEnSE Cohorts and University of Utah (N = 1,090)
Age at diagnosis, mean (s.d.)	55.9 ± 9.0	62.2 ± 11.4	54.4 ± 10.1
Average follow-up time, mean years (s.d.)	3.8 ± 2.5	5.3 ± 3.5	4.8 ± 3.3
Mortality, N (%)	103 (33.7)	174 (72.8)	423 (38.5%)
Type of disease (N with data)	292	239	649
IgG Kappa, N (%)	127 (43.5)	86 (36.0)	145 (22.3)
IgG Lambda, N (%)	44 (15.1)	49 (20.5)	64 (9.9)
IgG light chain not specified, N (%)	11 (3.8)	—	173 (26.7)
IgA Kappa, N (%)	31 (10.6)	35 (15.0)	35 (5.4)
IgA Lambda, N (%)	21 (7.2)	22 (9.1)	33 (5.1)
IgA light chain not specified, N (%)	—	—	75 (11.6)
Light chain only, N (%)	51 (17.5)	31 (13.0)	86 (13.3)
Other/nonsecretory, N (%)	7 (2.4)	16 (6.7)	38 (5.9)
ISS stage (N with data)	140	221	513
1, N (%)	49 (35.0)	86 (38.9)	162 (31.6)
2, N (%)	55 (39.3)	74 (33.5)	165 (32.2)
3, N (%)	36 (25.7)	61 (27.6)	186 (36.3)
Durie-Salmon stage (N with data)	256	187	633
IA or IB, N (%)	39 (15.2)	14 (7.5)	69 (10.9)
IIA or IIB, N (%)	40 (15.6)	30 (16.4)	149 (23.5)
IIIA or IIIB, N (%)	177 (69.2)	143 (76.5)	415 (65.6)

IMMEnSE, International Multiple Myeloma rESEarch; UCSF, University of California San Francisco.



**Figure 1 | GWAS results for survival among MM patients.** (a) Manhattan plot of results of genome-wide association analysis for survival in MM patients. Each point represents the negative log *P* value of the meta-analysis for association with survival using the UCSF and Mayo Clinic data. (b) Locuszoom plot for association statistics at the 16p13 region in the meta-analysis of UCSF and Mayo. Each dot represents the negative log *P* value for the association statistics from the meta-analysis. The top associated SNP (rs72773978) is coloured in purple and the remaining SNPs are coloured according to linkage disequilibrium values (*r*<sup>2</sup>) with the top SNP.

(Supplementary Table 2). We found that the effect was consistent among patients enrolled between 0 and 5.9 months after diagnosis (HR: 3.23; 95% CI: 1.28–8.17; *P* = 0.013; proportional

hazards model) those enrolled between 6 and 11.9 months after diagnosis (HR: 3.04; 95% CI: 1.67–5.55; *P* = 0.00028; proportional hazards model) and those enrolled from 12 to 23.9 months

(HR: 5.06; 95% CI: 1.18–21.79;  $P=0.029$ ; proportional hazards model). Thus, we concluded that the delay between diagnosis and ascertainment within the first 2 years was unlikely to affect the association between the SNP and overall survival. We also considered models that adjusted for the difference between date of diagnosis and ascertainment in the Mayo study and found no change in the association as expected (Supplementary Table 2), since nearly all of the participants were ascertained within 1 month of diagnosis.

We searched for additional SNPs that were associated independently of the top SNP by performing conditional analyses including rs72773978 and other SNPs within 100 kb of that SNP. We performed survival analyses for all SNPs that were either in no lethal dose (LD;  $R^2 < 0.1$ ) or in modest LD ( $R^2$  0.1–0.5) with RS72773978, adjusting for the effect of principal components (PC's) and for the effect of RS72773978. We found no other significant associations in the locus after adjusting for multiple hypothesis testing.

The median survival was decreased by  $\sim 2.7$  years among patients who were either heterozygous or homozygous for the rare variant (T allele) of rs72773978 compared with patients homozygous for the common variant in both the UCSF and Mayo Clinic cohorts (Fig. 2). In models that adjusted for age, gender and genetic ancestry,  $\sim 10$ –14% of patients had an increased risk of death (hazard ratio  $\sim 2.6$ ; Table 2) in the meta-analysis of these two data sets.

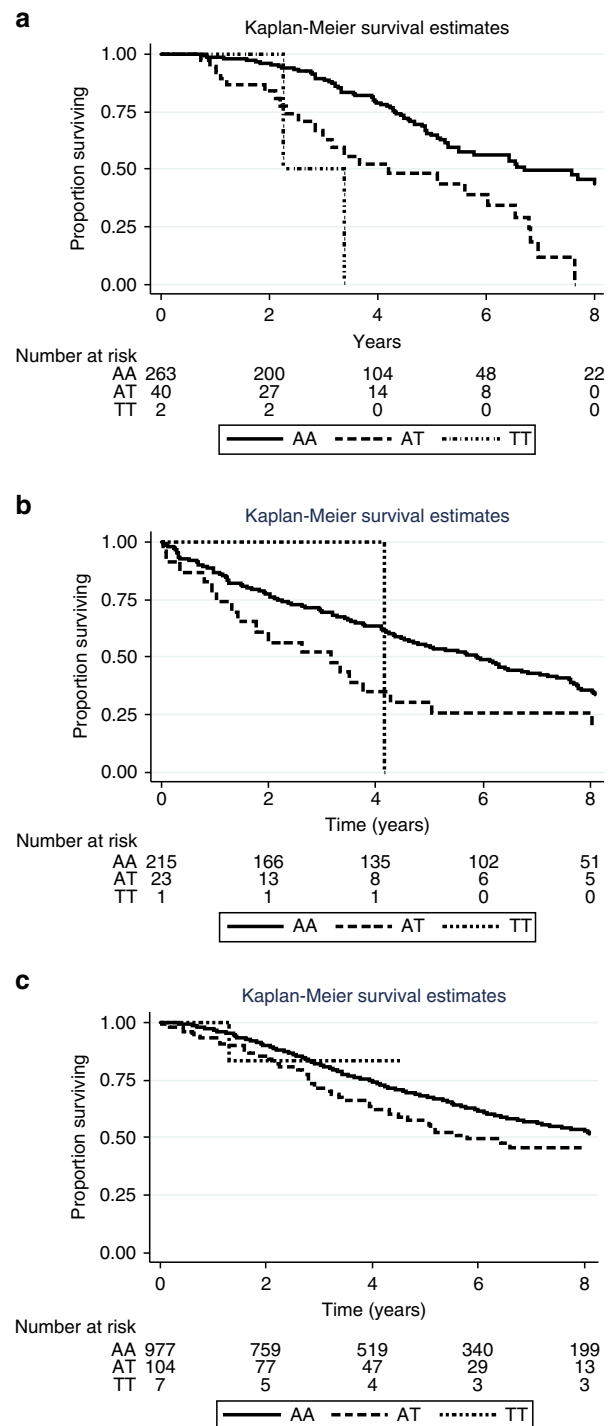
We also performed a separate analysis of individuals who genetically clustered with Caucasians in the UCSF data set (Supplementary Fig. 2). We repeated the meta-analysis of the Mayo Clinic data with the UCSF Caucasian-only sample (Supplementary Table 3), and found that the association with rs72773978 remained significant ( $P=2.4 \times 10^{-9}$ ; proportional hazards model). The UCSF cohort included African American patients ( $N=25$ ) and patients of Latino ( $N=27$ ) or other patients who clustered with those of mixed ancestry ( $N=24$ ). In an analysis of these patients, we also found a nominally significant association between shorter survival and the minor allele (HR: 2.43; 95% CI: 1.09–5.39;  $P=0.029$  proportional hazards model).

We also examined whether the top SNPs that we identified were associated with susceptibility to MM in our two studies. We found no significant difference in genotype frequencies between cases and controls (Supplementary Table 4).

**Relationship to stage at diagnosis and treatment.** We evaluated the effect of the genotype on survival after adjustment for the clinical stage. In analyses that adjusted for stage using either the international staging system (ISS) definition or the Durie–Salmon staging system, the rs72773978 genotype remained a strong predictor of survival (Table 2). Among the participants in the Mayo Clinic on whom LDH levels were available ( $N=154$ ), we saw a consistent level of association (HR: 2.25; 95% CI: 1.26–4.03;  $P=0.006$ ; proportional hazards model).

Since MM treatment has improved significantly in the last decade, we used data available from both cohorts to determine whether the SNP effect varied by initial treatment (Supplementary Table 5). We found a consistent effect of the SNP regardless of the type of treatment initiated (Table 3) and no evidence of interaction between treatment and the SNP in either the UCSF ( $P=0.9$ ;  $P$  for interaction derived using proportional hazards model) or Mayo Clinic ( $P=0.52$ ;  $P$  for interaction derived using proportional hazards model) cohorts.

Among patients in the Mayo Clinic study, 134 (56%) were treated with both high-dose chemotherapy (HDC) followed by autologous stem cell rescue. We adjusted for HDC in the proportional hazards models and found that, although HDC was



**Figure 2 | Kaplan-Meier Survivorship plot by genotype for rs72773978.** (a) UCSF ( $N=306$ ); (b) Mayo ( $N=239$ ) and (c) IMMENSE/Utah ( $N=1,080$ ) studies. We plotted the results by genotype. To generate the plots for the discovery data sets, we rounded the imputation results to the nearest whole numbers to infer genotypes of individuals.

a strong predictor of longer survival in the cohort, there was only mild attenuation of the SNP association with survival (Supplementary Table 6). Since nearly all (97%) of the UCSF patients received HDC, the same analysis could not be performed within the UCSF study.

**Replication.** We replicated the association of top SNPs from the UCSF and Mayo Clinic meta-analysis in a replication meta-

**Table 2 | Association of rs72773978 and MM survival.**

Study	Allele frequency	HR	95% CI	P value*
Age- and sex-adjusted <sup>†</sup>				
UCSF (n = 306)	0.075	3.03	1.95–4.73	8.0 × 10 <sup>-7</sup>
Mayo (n = 239)	0.054	2.31	1.51–3.53	1.1 × 10 <sup>-4</sup>
Discovery meta-analysis (UCSF and Mayo)		2.65	1.94–3.58	6.0 × 10 <sup>-10</sup>
Replication <sup>‡</sup> (1,090): IMMEnSE n = 772, Utah n = 315	0.051	1.34	1.01–1.78	0.044
Discovery and replication meta-analysis		1.93	1.54–2.41	6.7 × 10 <sup>-9</sup>
Age-, sex- and stage-adjusted <sup>†</sup>				
ISS-adjusted				
UCSF (n = 140)		4.06	2.04–8.05	6.3 × 10 <sup>-5</sup>
Mayo (n = 221)		2.41	1.51–3.83	2.2 × 10 <sup>-4</sup>
Replication <sup>‡</sup> (n = 513)		1.70	1.19–2.44	0.004
Salmon-Durie-adjusted				
UCSF (n = 256)		2.81	1.73–4.57	3.2 × 10 <sup>-5</sup>
Mayo (n = 184)		2.75	1.68–4.49	5.3 × 10 <sup>-5</sup>
Replication <sup>‡</sup> (n = 625)		1.46	1.04–2.05	0.029

CI, confidence interval; HR, hazard ratio; GWAS, genome-wide association study; PCA, principal components analysis; IMMEnSE, International Multiple Myeloma rESEarch; MM, multiple myeloma; UCSF, University of California San Francisco.

\*P values are calculated from proportional hazards models.

<sup>†</sup>The UCSF and Mayo Clinic study results are also adjusted with PCA. The IMMENSE data are not owing to the lack of availability of GWAS data.

<sup>‡</sup>The IMMEnSE results are from a meta-analysis of the individual regions (see Supplementary Table 3).

analysis of 1,090 MM cases, including 772 European MM patients from the IMMEnSE consortium and 318 from the Utah cohort (Supplementary Table 7). We selected two SNPs for replication including rs72773978, one of the top associated SNPs from the meta-analysis and rs12598966, a SNP resulting in an amino-acid substitution in the *FOPNL* gene. We performed a meta-analysis of all of the replication studies and found a significant association between rs72773978 and survival in the replication cohorts adjusted by age and sex (HR: 1.34; 95% CI: 1.01–1.74;  $P = 0.044$ ; proportional hazards model) with survival shorter by ~1.2 years among carriers of the minor allele (Fig. 2c). There was no evidence of heterogeneity of effect within the replication studies ( $P = 0.14$ ;  $\chi^2$ -test for heterogeneity). The other SNP, rs12598966, was not significantly associated with survival in the replication.

We noted a slightly stronger effect size in analyses that adjusted for stage. However, the change in effect size was not due to negative confounding between stage and the SNP, but rather to the fact that the cohorts with missing data on stage were the ones with an inconsistent point estimate (Supplementary Table 3). In a meta-analysis that did not adjust for the ISS stage but just included participants without missing data on the ISS stage, we saw approximately the same effect size (HR: 1.71; 95% CI: 1.18–2.47;  $P = 0.005$ ; proportional hazards model) as in the meta-analysis that adjusted for the ISS stage (Supplementary Table 7).

**Analysis of function.** The top SNPs were in a region that overlapped the entire *FOPNL* gene and a portion of the *MYH11* gene (Fig. 1b). In addition, known drug transporters, *ABCC1* and *ABCC6*, are located ~50 and 300 kb away, respectively. One of the top SNPs, rs12598966, is located in the coding sequence of *FOPNL* and leads to a nonsynonymous amino-acid substitution: Glu->Lys at amino acid 156; however, this SNP was not significantly associated with survival in the replication and was not predicted to have a deleterious effect on protein function (sorting intolerant from tolerant (SIFT) score = 0.89 and Polyphen 2

**Table 3 | Effect of SNP by initial treatment among patients in the UCSF cohort.**

	HR*	95% CI	P value <sup>†</sup>
Mayo Clinic			
Old treatments <sup>‡</sup>		N = 136, 102 deaths	
RS72773978	1.90	0.98–3.83	0.057
New treatments <sup>§</sup>		N = 93, 64 deaths	
RS72773978	2.71	1.56–4.70	0.00045
Entire sample adjusted for treatment, N = 229, 166 deaths			
RS72773978	2.18	1.43–3.32	0.00028
UCSF			
Old treatments <sup>‡</sup>		N = 124, 80 deaths	
RS72773978	3.37	1.76–6.50	0.00026
New treatments <sup>§</sup>		N = 208, 33 deaths	
RS72773978	3.62	1.73–7.56	0.0006
Entire sample adjusted for treatment			
RS72773978	3.35	2.07–5.41	8.2 × 10 <sup>-7</sup>

CI, confidence interval; HR, hazard ratio; GWAS, genome-wide association study; SNP, single-nucleotide polymorphism; UCSF, University of California San Francisco.

\*All models are adjusted for age, gender and principal components 1–10.

<sup>†</sup>P values are calculated from proportional hazards models.

<sup>‡</sup>Regimens including vincristine/adriamycin/dexamethasone or melphalan/prednisone.

<sup>§</sup>Treatments containing at least one of the following agents: thalidomide, bortezomib or lenalidomide.

score = 0.275). Next, we investigated the top 145 SNPs (all in tight LD ( $r^2 > 0.8$ ) with rs72773978) for an effect on gene expression using GENEVAR<sup>21</sup>. The top two SNPs, rs72773978 and rs117863986, were not included in the database; however, six other SNPs in strong LD ( $r^2 = 1$ ) with rs72773978 are present in the database and are associated with expression of *FOPNL* (Supplementary Table 8). The minor allele of these SNPs predicted higher expression of *FOPNL*. There was no significant association between these SNPs and expression of other genes within 1 MB of the locus.

We identified 13 SNPs in LD with rs72773978 as being potentially functional (Supplementary Table 9). Six of these SNPs are in the 3' untranslated repeat of the *FOPNL* gene and, therefore, may be involved in transcript stability. Seven SNPs were identified as being in sites of open chromatin and thus may be involved in transcriptional regulation.

Since the top SNPs were associated with gene expression, we hypothesized that expression of *FOPNL* may be associated with survival among MM patients. In particular, higher expression of *FOPNL* is associated with the minor allele of the top SNPs from the GWAS and should also be associated with shorter survival. We used publicly available data on gene expression (GSE2658) and survival from 414 MM cases to test this hypothesis<sup>13</sup>. As predicted, we found a significant association between higher expression of *FOPNL* and worse survival (Supplementary Table 10). *FOPNL* is known to localize to the centrosome and the pericentriolar satellites. Since centrosome amplification is known to be a predictor of poor prognosis, we evaluated the association between *FOPNL* expression and the centrosome index, a previously validated gene expression signature of centrosome amplification. We found a very strong correlation (Supplementary Table 11) between higher *FOPNL* expression and increased centrosome index in the study we analysed for survival (GSE2658) and in two additional studies (GSE19784 and GSE26760).

## Discussion

We performed a GWAS for survival among MM patients and identified SNPs at chromosome 16p13 that were strongly

associated with mortality. The SNPs were in the region of the *FOPNL* gene and a subset of the SNPs were associated with *FOPNL* expression levels, with the minor allele predicting higher expression<sup>22</sup>. We also found that *FOPNL* expression was associated with poorer survival using data from a previous study<sup>13</sup>. Thus, our results strongly suggest that *FOPNL* is a gene involved in myeloma progression.

*FOPNL* is known to be associated with centrosome function<sup>23,24</sup>. Centrosome amplification is common in MM and is associated with poor prognosis<sup>25,26</sup>. Furthermore, inhibition of centrosomal clustering may be effective in treatment of MM<sup>27,28</sup>. We found that germline variation that affects a gene involved in centrosomal function may also contribute to disease progression. Furthermore, we found an association between *FOPNL* expression and centrosome index in three data sets of gene expression from myeloma samples. However, our results implicating *FOPNL* as the causal gene at this locus rely on the synthesis of several data sets. We were not able to directly correlate the SNPs with gene expression, centrosome index and survival in the same data set. It is possible that another gene/s at this locus may be responsible for the effect we observe, or that the effect is mediated by *FOPNL*, but that it is not through a mechanism related to centrosome amplification. Additional studies of this gene and centrosomal function will help to further define the mechanism underlying the association that we identified.

Our results imply that germline genetic variation is associated with survival among patients with MM. Other GWAS have identified loci that affect survival in other cancer types<sup>15–18,20</sup>. At least one of the known loci for MM susceptibility is associated with the risk for a particular subtype of MM<sup>4</sup>, which may also have an effect on prognosis.

MM is a heterogeneous disease with substantial variation in prognosis among different patients. Identifying patients who are at higher risk of progression may be of importance in treating these patients more aggressively earlier in their disease. Our results identify the *FOPNL* genotype as a predictor of survival, and we found that the association remains significant after adjustment for stage at diagnosis. However, the effect size we observed in the replication cohorts was substantially attenuated compared with the effect size in the discovery cohorts. This difference is most likely because of ‘winner’s curse’—a tendency for the initial study to overestimate the true effect size<sup>29</sup>. Thus, the replication cohorts in our study are more likely to represent the true effect size in future studies.

Our study has several important limitations. We could not examine the association between SNPs and MM survival by cytogenetic abnormalities since the majority of our patients were diagnosed before the common use of cytogenetic and fluorescent *in situ* hybridization analyses in clinical practice. Therefore, it will be important to examine the utility of this genotype in the setting of cytogenetic and fluorescent *in situ* hybridization analyses and gene expression signatures. Furthermore, we could not adjust for gene expression patterns, which are also known to be associated with survival.

We found no difference between patients who were initially treated with modern versus older therapies. However, our ability to analyse the SNP using different therapies was limited to the UCSF data set and we had inadequate power to detect interactions between the SNPs and particular drugs. Furthermore, we only adjusted for the association between initial treatment and survival, and it is likely that many of the patients who were initially treated with older regimens received newer regimens if they survived to the era when these became available. It is possible that the effect that we saw is modified by one regimen or by one class of medications. Additional studies should

be conducted in the context of clinical trials or other cohorts where treatment regimens are more uniform to investigate whether the effect of the locus we identified is modified by particular treatments.

In summary, we found a strong association between a locus on chromosome 16p and MM survival that is likely due to an effect on expression of the *FOPNL* gene. The SNPs we identified may become important clinical predictors of outcome among MM patients.

## Methods

**Participants.** *UCSF study:* The UCSF Institutional Review Board (IRB) approved ascertainment of cases and use of existing biospecimens for genetic analysis. All participants gave informed consent. The study included 370 patients treated for MM at UCSF between 1989 and 2010. We excluded 10 samples because of insufficient clinical data. We also excluded 42 participants whose blood was collected  $\geq 2$  years after diagnosis from the survival analysis because of the potential bias towards long-term survivors among these participants. The median delay between diagnosis and ascertainment among the 42 participants excluded was 2.9 years (interquartile range 2.4–4.7 years). The median delay among the participants included was 7.6 months (interquartile range 5.7–8.9 months).

We used white blood cells harvested after mobilization of stem cells with granulocyte colony-stimulating factor as a source of DNA. Mobilization of stem cells is performed before HDC followed by autologous stem cell transplantation. The patient receives granulocyte colony-stimulating factor and then undergoes harvesting of peripheral white blood cells via apheresis several days later<sup>30,31</sup>. Bone marrow stem cell fraction is monitored via CD34 antibody, and apheresis is continued until an adequate number of CD34+ cells have been collected for stem cell rescue. The cells are then stored in liquid nitrogen.

We ascertained the date of death using chart reviews and death registry data. Survival time was determined as the date of diagnosis until date of death or last clinic visit for patients who were not known to have died. Clinical stage and initial chemotherapy regimen was determined by chart review. For analysis of treatment, we dichotomized treatments into either newer regimens (including an Imid and/or proteasome inhibitor) or older regimens (including neither an Imid nor proteasome inhibitor). Nearly all of the participants in the UCSF study (>97%) received HDC with stem cell rescue (autologous bone marrow transplant).

*Mayo Clinic study:* Ascertainment of MM cases and genotyping was approved by the Mayo Clinic IRB. The study included incident MM cases seen in the regional practice between 1998 and 2007 and recruited within 6 months of initial diagnosis. Nearly all participants (96%) were recruited within 1 month of diagnosis and, the remainder were recruited between 1 and 5.5 months after diagnosis. Eligible cases provided consent and a blood sample for research studies of MM. A total of 243 MM cases were used for analyses. DNA was extracted from stored peripheral blood samples. We ascertained date of death and calculated survival time as described for the UCSF study. The clinical stage was determined by chart review. Clinical data on survival could be ascertained on 239 patients. Initial chemotherapy regimen and use of HDC was determined by chart review. For analysis of treatment, we dichotomized treatments into either newer regimens (including an Imid and/or proteasome inhibitor) or older regimens (including neither an Imid nor proteasome inhibitor). Approximately 56% of participants were treated with autologous stem cell transplant.

*IMMENSE study:* The IMMENSE consortium is a case-control study recruited from seven different European and North American countries<sup>32</sup>. MM cases are defined by a confirmed diagnosis of MM, according to the International Myeloma Working Group criteria<sup>33</sup>. For each patient, demographic and clinical parameters at diagnosis and survival were collected by the responsible clinicians in each of the IMMENSE centres. The data collected are standardized in a central database kept at the German Cancer Research Center (DKFZ, Heidelberg, Germany). For each subject, a sample of peripheral blood or extracted DNA has been collected and sent to DKFZ. A total of 772 MM cases with survival information available in the IMMENSE consortium were included in this study.

*Utah study:* Sampling and genetic analysis was approved by the University of Utah IRB. The study included prevalent MM cases in the state of Utah, ascertained up to 2012. Eligible cases provided consent and a blood or saliva sample from which DNA was extracted. Date of diagnosis was confirmed from chart review and Utah Cancer Registry data. Date of death was confirmed from chart review and death registry data. Survival time was determined as the date of diagnosis until the date of death, last contact with the study or last known event in Utah (determined from statewide vital records, driver’s license renewals and voter registrations in the Utah Population Database) for patients who were not known to have died. A total of 318 MM cases with DNA and survival information were available for this study.

**GWAS genotyping and imputation.** All SNP positions were annotated using the Genome Reference Consortium GRCh37 (Hg19) version of the human genome.

*UCSF:* A pilot study of 81 MM samples used an Illumina 660 array genotyped at the UCSF Genomics Core Facility. In a second phase, we genotyped 289 MM samples using an Illumina Omni5 array at Expression Analysis (Durham, NC).

Of the 370 participants in the GWAS, 52 participants were excluded from the survival analysis as noted above because of either insufficient clinical data ( $N = 10$ ) or because of  $> 2$  year time difference between diagnosis and ascertainment ( $N = 42$ ). Of the remaining 318 samples we dropped 12 since they did not pass quality control for genotyping. Eleven were dropped because of high missing genotype values ( $> 5\%$  missing genotypes per sample) and one sample was dropped because of potential contamination, leaving 306 patients. We dropped SNPs that had  $> 5\%$  missing values or were monomorphic. Imputation was performed using IMPUTE2 (ref. 34; [https://mathgen.stats.ox.ac.uk/impute/impute\\_v2.html#home](https://mathgen.stats.ox.ac.uk/impute/impute_v2.html#home)) with all samples from 1000 Genomes data set (Version 2, May 2011 release<sup>35</sup>) as a reference. Imputed SNPs with Information  $< 0.5$  or  $MAF < 0.025$  were excluded, leaving 8,036,255 SNPs for analysis.

**Mayo Clinic:** Cases were genotyped using the Affymetrix 6.0 array. Monomorphic SNPs and those with a call rate  $< 95\%$  were excluded, leaving 786,950 observed SNPs. Four samples with call rates  $< 95\%$  and one sample with non-European ancestry according to principal components analysis were excluded, leaving 243 MM cases and 239 with follow-up past date of diagnosis. Imputation was performed with BEAGLE<sup>36</sup> (<http://faculty.washington.edu/browning/beagle/beagle.html>), using all samples from version 2 of the 1,000 Genomes data (May 2011 release) as reference. Imputed SNPs with an  $r^2 < 0.3$  or  $MAF < 0.025$  were excluded, leaving 7,276,170 SNPs for further analyses.

**Replication genotyping.** We selected two SNPs for replication in IMMENSE and Utah samples including one of the top two SNPs from the meta-analysis (rs72773978) and a SNP in high linkage disequilibrium with the top SNP, encoding a nonsynonymous amino-acid substitution in *FOPNL* (rs12598966). These SNPs were typed using 5' exonuclease (TaqMan) assays (ABI) at the German Cancer Research Center (DKFZ) in Heidelberg (IMMENSE samples) and at the Genomics Core at the University of Utah (Utah samples). Duplicates of 12% of the samples were interspersed throughout the plates and concordance rate among duplicates was  $> 99.9\%$ .

**Statistical analysis.** We performed genome-wide analyses for association with survival using proportional hazards models in the UCSF and Mayo data sets separately. We inferred genetic ancestry using principal components analysis (PCA) in each cohort using SmartPCA<sup>37</sup>. Each SNP was entered into the model under an assumption of log-additive increased risk, and adjusting for PC's 1–3, age and gender. Imputed SNPs were modelled using the probability of genotypes. We tested the proportional hazards assumption for the top SNPs by calculating the scaled Schoenfeld residuals and testing whether they are significantly associated with time<sup>38</sup>.

We also performed a subset analysis of Caucasians only in the UCSF data set. We identified Caucasians on the basis of genetic ancestry (see Supplementary Fig. 2). Individuals who clustered with self-described Caucasians ( $PC1 > 0$ ,  $PC2 < 0$ ) were included in this subset analysis ( $N = 229$ ).

All analyses were performed in R. For graphing survival results, we used the Kaplan–Meier estimates of the survival function and graphed the results using Stata (Version 10). We graphed the association statistics for all SNPs near the top locus using LocusZoom<sup>39</sup>.

We used Cox regression models adjusted for age and gender to test the association between SNPs and survival in the IMMENSE consortium and the Utah cohort.

We performed a meta-analysis of the UCSF and Mayo Clinic results on a total of 6,026,834 SNPs in common from both GWAS that met the allele frequency and imputation quality thresholds. We also conducted a meta-analysis of data on two top SNPs from seven centres within the IMMENSE consortium and the Utah study in our replication study. We calculated a fixed effects model for each SNP using METAL<sup>40</sup>. We used Cochran's Q statistic to test for heterogeneity.

To examine the association of SNPs and risk of MM, we compared the genotype frequencies of cases versus ethnically matched controls from the UCSF ( $N = 298$ ) and the Mayo Clinic ( $N = 295$ ) sites, respectively. We used logistic regression models, adjusting for PC1–3 age and gender.

**Expression quantitative trait locus (eQTL) analysis.** We used the data set from ref. 22 for eQTL analyses, which consists of 856 Caucasian individuals including 154 monozygotic twin pairs, 232 dizygotic twin pairs and 84 singletons. We focused on expression in lymphocytes in this data set. We used GENEVAR<sup>21</sup> to query the top 145 SNPs from the GWAS and identified 6 SNPs, that were also in the data set described in ref. 22. We queried GENEVAR for beta coefficients and P values for associations between six SNPs and the genes within a 1-Mb window, including *FOPNL*, *MYH11*, *ABCC1*, *NDE1*, *KIAA0430*, *ABCC6*, *RRN3*, *NTAN1*, *KIAA0250* and *KIAA0251*.

**Gene expression and centrosome index.** We downloaded gene expression data from refs 13,41,42 from the National Institutes of Health Gene Expression Omnibus (accession number: GSE2658, GSE19784 and GSE26760, respectively). Reference 13 consisted of gene expression data from 559 MM samples assayed on Affymetrix U133 arrays; ref. 41 consisted of gene expression data from purified CD138+ plasma cells of 320 newly diagnosed myeloma patients using Affymetrix

GeneChip U133 plus 2.0 arrays; ref. 42 consisted of 304 CD138-purified bone marrow samples from patients with MM were analysed on Affymetrix U133 Plus 2.0 microarrays.

**Gene expression and MM outcome:** Of the samples in the Zhan *et al.*<sup>13</sup> data set, 414 also had available clinical data and were included in the original publication and, therefore, we used the data from these 414 samples in our analyses. We used log-transformed probe intensity values as predictors of survival, entering these as continuous variables into a proportional hazards model. We analysed each of the two probes for *FOPNL* on the Affymetrix U133 array separately and also considered the average of the two probes as a predictor of overall survival in the proportional hazards model.

**Analysis of potential SNP function:** We used SIFT<sup>43</sup> and Polyphen2 (ref. 44) to determine the likelihood that a nonsynonymous amino-acid substitution has a deleterious effect on protein function. We used FunciSNP<sup>45</sup> to determine whether any of the SNPs may affect gene expression, including any SNPs with  $R^2 > 0.7$  with rs72773928.  $R^2$  values for linkage disequilibrium were calculated in European ancestry samples from 1,000 genomes.

## References

- Howlader, N. *et al.* *SEER Cancer Statistics Review, 1975–2010* (National Cancer Institute, 2010).
- Greenberg, A. J., Rajkumar, S. V. & Vachon, C. M. Familial monoclonal gammopathy of undetermined significance and multiple myeloma: epidemiology, risk factors, and biological characteristics. *Blood* **119**, 5359–5366 (2012).
- 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).
- Weinhold, N. *et al.* The CCND1 c.870G > A polymorphism is a risk factor for t(11;14)(q13;q32) multiple myeloma. *Nat. Genet.* **45**, 522–525 (2013).
- Broderick, P. *et al.* Common variation at 3p22.1 and 7p15.3 influences multiple myeloma risk. *Nat. Genet.* **44**, 58–61 (2012).
- Rajkumar, S. V. & Greipp, P. R. Prognostic factors in multiple myeloma. *Hematology* **13**, 1295–1314 (1999).
- Kumar, S. *et al.* Trisomies in multiple myeloma: impact on survival in patients with high-risk cytogenetics. *Blood* **119**, 2100–2105 (2012).
- Fonseca, R. *et al.* Prognostic value of chromosome 1q21 gain by fluorescent *in situ* hybridization and increase CKS1B expression in myeloma. *Leukemia* **20**, 2034–2040 (2006).
- Fonseca, R. *et al.* Clinical and biologic implications of recurrent genomic aberrations in myeloma. *Blood* **101**, 4569–4575 (2003).
- Zhan, F. *et al.* Gene-expression signature of benign monoclonal gammopathy evident in multiple myeloma is linked to good prognosis. *Blood* **109**, 1692–1700 (2007).
- Zhan, F., Barlogie, B., Mulligan, G., Shaughnessy, J. D. Jr., & Bryant, B. High-risk myeloma: a gene expression based risk-stratification model for newly diagnosed multiple myeloma treated with high-dose therapy is predictive of outcome in relapsed disease treated with single-agent bortezomib or high-dose dexamethasone. *Blood* **111**, 968–969 (2008).
- Zhan, F. *et al.* Global gene expression profiling of multiple myeloma, monoclonal gammopathy of undetermined significance, and normal bone marrow plasma cells. *Blood* **99**, 1745–1757 (2002).
- Zhan, F. *et al.* The molecular classification of multiple myeloma. *Blood* **108**, 2020–2028 (2006).
- Zhan, F. *et al.* Gene expression profiling of human plasma cell differentiation and classification of multiple myeloma based on similarities to distinct stages of late-stage B-cell development. *Blood* **101**, 1128–1140 (2003).
- Wu, C. *et al.* Genome-wide association study identifies common variants in SLC39A6 associated with length of survival in esophageal squamous-cell carcinoma. *Nat. Genet.* **45**, 632–638 (2013).
- Kiyotani, K. *et al.* A genome-wide association study identifies locus at 10q22 associated with clinical outcomes of adjuvant tamoxifen therapy for breast cancer patients in Japanese. *Hum. Mol. Genet.* **21**, 1665–1672 (2012).
- Shu, X. O. *et al.* Novel genetic markers of breast cancer survival identified by a genome-wide association study. *Cancer Res.* **72**, 1182–1189 (2012).
- Innocenti, F. *et al.* A genome-wide association study of overall survival in pancreatic cancer patients treated with gemcitabine in CALGB 80303. *Clin. Cancer Res.* **18**, 577–584 (2012).
- Wu, C. *et al.* Genome-wide association study of survival in patients with pancreatic adenocarcinoma. *Gut* **63**, 152–160 (2012).
- Wu, C. *et al.* Genome-wide interrogation identifies YAP1 variants associated with survival of small-cell lung cancer patients. *Cancer Res.* **70**, 9721–9729 (2010).
- Yang, T. P. *et al.* Genevar: a database and Java application for the analysis and visualization of SNP-gene associations in eQTL studies. *Bioinformatics* **26**, 2474–2476 (2010).
- Grundberg, E. *et al.* Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nat. Genet.* **44**, 1084–1089 (2012).

23. Aubusson-Fleury, A. *et al.* The conserved centrosomal protein FOR20 is required for assembly of the transition zone and basal body docking at the cell surface. *J. Cell Sci.* **125**, 4395–4404 (2012).
24. Sedjai, F. *et al.* Control of ciliogenesis by FOR20, a novel centrosome and pericentriolar satellite protein. *J. Cell Sci.* **123**, 2391–2401 (2010).
25. Chng, W. J. *et al.* The centrosome index is a powerful prognostic marker in myeloma and identifies a cohort of patients that might benefit from aurora kinase inhibition. *Blood* **111**, 1603–1609 (2008).
26. Chng, W. J. *et al.* Clinical implication of centrosome amplification in plasma cell neoplasm. *Blood* **107**, 3669–3675 (2006).
27. Raab, M. S. *et al.* GF-15, a novel inhibitor of centrosomal clustering, suppresses tumor cell growth in vitro and in vivo. *Cancer Res.* **72**, 5374–5385 (2012).
28. Shiheido, H. *et al.* A phthalimide derivative that inhibits centrosomal clustering is effective on multiple myeloma. *PLoS ONE* **7**, e38878 (2012).
29. Zollner, S. & Pritchard, J. K. Overcoming the winner's curse: estimating penetrance parameters from case-control data. *Am. J. Hum. Genet.* **80**, 605–615 (2007).
30. Stewart, D. A. *et al.* Factors predicting engraftment of autologous blood stem cells: CD34+ subsets inferior to the total CD34+ cell dose. *Bone Marrow Transplant.* **23**, 1237–1243 (1999).
31. Stewart, D. A. *et al.* The CD34+90+ cell dose does not predict early engraftment of autologous blood stem cells as well as the total CD34+ cell dose. *Bone Marrow Transplant.* **25**, 435–440 (2000).
32. Martino, A. *et al.* Genetics and molecular epidemiology of multiple myeloma: the rationale for the IMMENSE consortium (review). *Int. J. Oncol.* **40**, 625–638 (2012).
33. International Myeloma Working Group. Criteria for the classification of monoclonal gammopathies, multiple myeloma and related disorders: a report of the International Myeloma Working Group. *Br. J. Haematol.* **121**, 749–757 (2003).
34. Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
35. Genomes Project C *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
36. Browning, B. L. & Browning, S. R. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am. J. Hum. Genet.* **84**, 210–223 (2009).
37. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2**, e190 (2006).
38. Grambsch, P. A. & Therneau, T. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* **81**, 515–526 (1994).
39. Pruim, R. J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336–2337 (2010).
40. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
41. Broyl, A. *et al.* Gene expression profiling for molecular classification of multiple myeloma in newly diagnosed patients. *Blood* **116**, 2543–2553 (2010).
42. Chapman, M. A. *et al.* Initial genome sequencing and analysis of multiple myeloma. *Nature* **471**, 467–472 (2011).
43. Sim, N. L. *et al.* SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res.* **40**, W452–W457 (2012).
44. Adzhubei, I., Jordan, D. M. & Sunyaev, S. R. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr. Protoc. Hum. Genet.* Chapter 7, Unit 7–20 (2013).
45. Coetzee, S. G., Rhie, S. K., Berman, B. P., Coetzee, G. A. & Noushmehr, H. FunciSNP: an R/bioconductor tool integrating functional non-coding data sets with genetic association studies to identify candidate regulatory SNPs. *Nucleic Acids Res.* **40**, e139 (2012).

## Acknowledgements

This work was supported by the Steve and Nancy Grand Multiple Myeloma Translational Initiative and a grant from Expression Analysis and a grant from the National Cancer Institute (R21CA191896). E.Z. is supported in part by a mid-career award in patient research from the National Cancer Institute (K24169004). Collection of blood samples from Polish patients and controls from Łódź area, and DNA extraction was supported by a grant from Polish Ministry of Science and Higher Education (No. NN402178334). DNA extraction from Danish healthy controls was supported by The Research Fund at Region Sjælland, Denmark. The Utah study was supported by LLS 6067–09, CA152336 and CA134674, with data collection made possible, in part, by the Utah Population Database (UPDB) and the Utah Cancer Registry (UCR). Partial support for all data sets within the UPDB is provided by the Huntsman Cancer Institute (HCI) and the HCI Cancer Center Support grant, P30 CA42014. The UCR is supported in part by NIH contract HHSN261201000026C from the NCI SEER Program with additional support from the Utah State Department of Health and the University of Utah. In Utah, we thank Jathine Wong for support in formatting files and advice on R programming. Several authors (E.Z., N.C., F.C., C.V.) are members of the International Multiple Myeloma Consortium. We are grateful to the leadership and other members of the International Multiple Myeloma Consortium for facilitating this collaborative study.

## Author contributions

Study conception: E.Z., E.D. and C.V.; collection of the data: E.Z., E.D., K.C., D.C., P.B., J.C.-J., C.D., M.D., A.G., K.J., A.J., S.K., D.A., H.M., T.M., J.M.-L., V.R., J.S., A.J.V., M.W., J.W., N.J.C., F.C. and C.V.; performed genotyping: A.M., R.D., A.L. and B.J.; analysis of the data: D.H., D.S., D.C., B.A. and Y.Z.; drafting of the manuscript: E.Z. and E.D.; critical revision of the manuscript: E.Z., E.D., D.H., A.M., D.S., K.C., D.C., B.A., P.B., G.B., Y.Z., J.C.-J., R.D., C.D., M.D., L.F., A.G., S.H., K.J., A.J., S.K., D.A., M.G., L.A.C.-A., B.J., A.L., H.M., T.M., J.M.-L., V.R., J.S., A.J.V., M.W., J.W., S.S., N.J.C., F.C. and C.V.

## Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

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

**How to cite this article:** Ziv, E. *et al.* Genome-wide association study identifies variants at 16p13 associated with survival in multiple myeloma patients. *Nat. Commun.* 6:7539 doi: 10.1038/ncomms8539 (2015).



# Corrigendum: Genome-wide association study identifies variants at 16p13 associated with survival in multiple myeloma patients

Elad Ziv, Eric Dean, Donglei Hu, Alessandro Martino, Daniel Serie, Karen Curtin, Daniele Campa, Blake Aftab, Paige Bracci, Gabriele Buda, Yi Zhao, Jennifer Caswell-Jin, Robert Diasio, Charles Dumontet, Marek Dudziński, Laura Fejerman, Alexandra Greenberg, Scott Huntsman, Krzysztof Jamrozik, Artur Jurczynszyn, Shaji Kumar, Djordje Atanackovic, Martha Glenn, Lisa A. Cannon-Albright, Brandt Jones, Adam Lee, Herlander Marques, Thomas Martin, Joaquin Martinez-Lopez, Vincent Rajkumar, Juan Sainz, Annette Juul Vangsted, Marzena Watek, Jeffrey Wolf, Susan Slager, Nicola J. Camp, Federico Canzian & Celine Vachon

*Nature Communications* 6:7539 doi: 10.1038/ncomms8539 (2015); Published 22 Jul 2015; Updated 9 Dec 2015

In this Article, members of the UCSF cohort who had been alive for longer than two years were inadvertently included in the data presented in Table 3. USCF/old treatments should have 109 patients with a hazard ratio of 3.35 and a *P* value of 0.00028 instead of the 124 patients with a hazard ratio of 3.37 and a *P* value of 0.00026. The USCF/new patients should have 187 patients with a hazard ratio of 3.57 and a *P* value of 0.0007 instead of the 208 patients with a hazard ratio of 3.62 and a *P* value of 0.0006. Finally, in the table legend, the first line should read ‘All models are adjusted for age, gender and principal components 1–3’. The exclusion of these individuals does not change the conclusions of the study. The correct version of Table 3 appears below.

**Table 3 | Effect of SNP by initial treatment among patients in the UCSF cohort.**

	HR*	95% CI	P value <sup>†</sup>
<i>Mayo Clinic</i>			
Old treatments <sup>‡</sup>		<i>N</i> = 136, 102 deaths	
RS72773978	1.90	0.98–3.83	0.057
New treatments <sup>§</sup>		<i>N</i> = 93, 64 deaths	
RS72773978	2.71	1.56–4.70	0.00045
Entire sample adjusted for treatment		<i>N</i> = 229, 166 deaths	
RS72773978	2.18	1.43–3.32	0.00028
<i>UCSF</i>			
Old treatments <sup>‡</sup>		<i>N</i> = 109, 60 deaths	
RS72773978	3.35	1.74–6.44	0.00028
New treatments <sup>§</sup>		<i>N</i> = 187, 30 deaths	
RS72773978	3.57	1.71–7.43	0.0007
Entire sample adjusted for treatment		<i>N</i> = 296, 90 deaths	
RS72773978	3.35	2.07–5.41	8.2 × 10 <sup>-7</sup>

\*All models are adjusted for age, gender and principal components 1–3.

<sup>†</sup>*P* values are calculated from proportional hazards models.

<sup>‡</sup>Treatments containing at least one of the following agents: thalidomide, bortezomib or lenalidomide.

<sup>§</sup>Regimens including vincristine/adriamycin/dexamethasone or melphalan/prednisone.