LETTERS

Identification of *RPS14* as a 5q$^-$ syndrome gene by RNA interference screen

Benjamin L. Ebert[1,2,3], Jennifer Pretz[1], Jocelyn Bosco[1], Cindy Y. Chang[1], Pablo Tamayo[1], Naomi Galili[4], Azra Raza[4], David E. Root[1], Eyal Attar[5], Steven R. Ellis[6] & Todd R. Golub[1,2,7]

Somatic chromosomal deletions in cancer are thought to indicate the location of tumour suppressor genes, by which a complete loss of gene function occurs through biallelic deletion, point mutation or epigenetic silencing, thus fulfilling Knudson's two-hit hypothesis[1]. In many recurrent deletions, however, such biallelic inactivation has not been found. One prominent example is the 5q$^-$ syndrome, a subtype of myelodysplastic syndrome characterized by a defect in erythroid differentiation[2]. Here we describe an RNA-mediated interference (RNAi)-based approach to discovery of the 5q$^-$ disease gene. We found that partial loss of function of the ribosomal subunit protein RPS14 phenocopies the disease in normal haematopoietic progenitor cells, and also that forced expression of RPS14 rescues the disease phenotype in patient-derived bone marrow cells. In addition, we identified a block in the processing of pre-ribosomal RNA in RPS14-deficient cells that is functionally equivalent to the defect in Diamond–Blackfan anaemia, linking the molecular pathophysiology of the 5q$^-$ syndrome to a congenital syndrome causing bone marrow failure. These results indicate that the 5q$^-$ syndrome is caused by a defect in ribosomal protein function and suggest that RNAi screening is an effective strategy for identifying causal haploinsufficiency disease genes.

The 5q$^-$ syndrome was reported in 1974 as the first chromosomal deletion in cancer associated with a distinct clinical phenotype[2]. Patients have a severe macrocytic anaemia, normal or elevated platelet counts, normal or reduced neutrophil counts, erythroid hypoplasia in the bone marrow, and hypolobated micromegakaryocytes[3]. These patients also have a propensity to progress to acute myeloid leukaemia (AML), although more slowly than other forms of myelodysplastic syndrome (MDS)[4]. A main cause of morbidity and mortality for these patients is the erythroid defect, which often requires continuing transfusions of red blood cells resulting in iron overload and subsequent organ dysfunction[4]. The 5q$^-$ syndrome is also unique because this subtype of MDS shows a remarkable response to treatment with the thalidomide analogue lenalidomide, although the mechanism of action of lenalidomide remains unknown[5].

Over the past 30 years, physical mapping methods have been used to narrow the region of recurrent somatic deletion on 5q to a 1.5-megabase common deleted region (CDR) containing 40 genes[6]. No patients with the 5q$^-$ syndrome have been reported to have biallelic deletions within the CDR, and no point mutations have been reported in the remaining allele of any of the 40 genes in the region. This observation led us to speculate that the 5q$^-$ syndrome may be caused by haploinsufficiency, suggesting that an alternative approach would be required to identify the gene responsible. We therefore examined whether the principal hallmarks of the disease (an erythroid maturation block with preservation of megakaryocyte differentiation) could be recapitulated experimentally with short hairpin RNAs (shRNAs) targeting each of the genes within the CDR.

We designed multiple lentivirally expressed shRNAs for each of the candidate genes, to control for possible off-target effects of any individual shRNA. The shRNAs were introduced into normal CD34$^+$ human haematopoietic progenitor cells, and the cells were induced to differentiate for 10 days along the erythroid and megakaryocytic lineages. The effect of each shRNA was assessed by fluorescence-activated cell sorting (FACS) analysis with the use of erythroid-specific and megakaryocyte-specific cell surface markers. The shRNAs targeting one gene, *RPS14*, recapitulated the 5q$^-$ syndrome phenotype: a severe decrease in the production of erythroid cells with relative preservation of megakaryocytic cells (Fig. 1). Furthermore, using the sequential expression of CD71 and glycophorin A during erythroid differentiation (Supplementary Fig. 1), we found that shRNAs targeting *RPS14* blocked the production of terminally differentiated erythroid cells, which is also consistent with the 5q$^-$ syndrome disease phenotype (Supplementary Fig. 2). In a statistical analysis that grouped all shRNAs targeting each gene into a single set, *RPS14* was the only gene that significantly altered differentiation (Supplementary Fig. 3). On the basis of these results, we focused our attention on *RPS14* as a candidate disease gene.

We first confirmed that all five *RPS14* shRNAs that scored in the screen in fact knocked down *RPS14* expression, and that the level of protein expression was of the order of half of the luciferase control cells, which is consistent with a model of *RPS14* haploinsufficiency (Fig. 2a). Each of the *RPS14* shRNAs decreased erythroid differentiation relative to megakaryocytic differentiation (Fig. 2b) and also caused a mild defect in erythroid versus myeloid differentiation (Fig. 2c), precisely as seen in patients with the clinical syndrome. *RPS14* knockdown also caused an increase in the ratio of immature-to-mature erythroid cells (Fig. 2d), as well as increased apoptosis of differentiating erythroid cells (Fig. 2e), which is consistent with the well-described apoptotic phenotype of MDS[7]. Given the possibility that multiple genes in the CDR might act in collaboration[8], we tested whether other effective shRNAs might increase the effect of *RPS14* knockdown. None of the combinations was more effective than *RPS14* shRNAs alone, suggesting that *RPS14* is the critical gene in the region that explains the haematopoietic differentiation defect associated with 5q$^-$ syndrome (Supplementary Fig. 4).

To confirm that RPS14 deficiency truly affects the erythroid differentiation programme (rather than simply modulating the expression of specific FACS markers), we performed genome-wide expression profiling of cells infected with control or *RPS14* shRNAs.

[1]Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA. [2]Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts 02115, USA. [3]Division of Hematology, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. [4]Division of Hematology Oncology, University of Massachusetts Medical School, Worcester, Massachusetts 01655, USA. [5]Division of Hematology/Oncology, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [6]Department of Biochemistry and Molecular Biology, University of Louisville, Kentucky 40292, USA. [7]Howard Hughes Medical Institute, Chevy Chase, Maryland 20815, USA.

We used gene set enrichment analysis (GSEA)[9] to assess the effect of *RPS14* knockdown on experimentally derived signatures of erythroid and megakaryocytic differentiation (Fig. 2f). As expected, the gene expression pattern of *RPS14* knockdown showed a significant abrogation of the erythroid differentiation signature ($P < 0.001$; Fig. 2g), in the setting of increased signature of neutrophil and platelet differentiation ($P < 0.001$ for both; Fig. 2h, i). In addition, *RPS14* shRNAs induced a signature of sensitivity to lenalidomide, the only drug approved by the Food and Drug Administration specifically for MDS patients with 5q deletions[5] (Supplementary Fig. 5). The *RPS14* shRNAs knocked down *RPS14* expression by an average of about 60% in these samples, which is consistent with haploinsufficiency as the cause of these phenotypes. To exclude further the possibility of biallelic inactivation of *RPS14*, we sequenced the *RPS14* gene in 32 MDS patient samples and subjected a subset of these samples to high-density single-nucleotide-polymorphism-based copy number analysis and gene expression profiling. In no case did we detect *RPS14* point mutations, cryptic biallelic deletions or loss of expression (for example, by aberrant methylation; see Supplementary Fig. 6). Taken together, these experiments show that partial loss of function of *RPS14* recapitulates the phenotype of the 5q$^-$ syndrome.

RPS14 is a component of the 40S ribosomal subunit, but the function of RPS14 in human cells has not been defined. To determine the effect of partial loss of function of RPS14 on pre-rRNA processing, we performed northern blotting of rRNA transcripts and sucrose-gradient analysis of intact polysomes. Decreased expression of RPS14 resulted in an accumulation of the 30S pre-rRNA species with a concomitant decrease in levels of 18S/18SE rRNA levels (Fig. 3a, b), which is consistent with reports for *Saccharomyces cerevisiae* that RPS14 is required for the processing of 18S pre-rRNA[10]. Specifically, a fourfold to ninefold increase in the 30S/18SE ratio was observed in cells expressing *RPS14* shRNAs. In addition, *RPS14* knockdown abrogated formation of the 40S subunit (Fig. 3c and Supplementary Fig. 7). The increased 30S/18SE ratio in RPS14-deficient cells was not simply a consequence of cell death: the ribosomal processing defect occurred before the onset of significant apoptosis, and pharmacologically induced apoptosis failed to generate the characteristic 30S/18S defect (Supplementary Figs 8 and 9). These results indicate that the block in pre-rRNA processing is a specific consequence of RPS14 deficiency.

An increase in the 30S/18SE ratio was observed in bone marrow cells from patients with 5q$^-$ syndrome in comparison with those from normal marrow (Fig. 3d), suggesting that a pre-rRNA processing defect does indeed occur in cells from patients. We note that the samples from patients contain a mixture of normal and 5q$^-$ disease cells, probably explaining why the 30S/18SE ratio is less perturbed than that seen in the experimental setting. The essential nature of RPS14 in ribosome biogenesis also probably explains why a complete loss of RPS14 (for example, through biallelic deletions) is never seen in cells from patients with 5q$^-$ syndrome. Complete loss of RPS14 is probably incompatible with cell survival, as it is in yeast[10].

To establish further that RPS14 deficiency accounts for the haematopoietic defect characteristic of 5q$^-$ syndrome, we attempted to rescue the erythroid differentiation defect in patient-derived bone marrow cells by using an RPS14 expression construct. CD34$^+$ cells from viably frozen bone marrow mononuclear cells obtained from MDS patients with and without 5q deletions (Supplementary Table 4) were induced to undergo differentiation *in vitro*. FACS analysis showed that in comparison with control, lentiviral expression of *RPS14* increased erythroid differentiation in patients with the 5q$^-$ syndrome but failed to do so in patients lacking 5q deletions ($P = 0.004$ for erythroid relative to megakaryocytic differentiation; $P = 0.0003$ for erythroid relative to myeloid; Fig. 4 and Supplementary Fig. 10). Furthermore, gene expression profiling coupled with GSEA showed that ectopic expression of RPS14 induced the gene expression signature of erythroid differentiation in 5q$^-$ syndrome patient samples ($P < 0.001$; Supplementary Fig. 11). These data demonstrate that overexpression of RPS14 rescues the erythroid differentiation defect seen in patients with 5q$^-$ syndrome, and establishes *RPS14* as the likely disease-causing gene.

Loss of function of a ribosomal protein might at first seem like an unlikely explanation for a disease with such a distinct haematopoietic phenotype. However, germline heterozygous mutations for two other ribosomal proteins—RPS19 and RPS24—have recently been described in the congenital disorder known as Diamond–Blackfan anaemia[11,12]. The phenotype of Diamond–Blackfan anaemia is strikingly similar to the 5q$^-$ syndrome: patients have a severe anaemia,

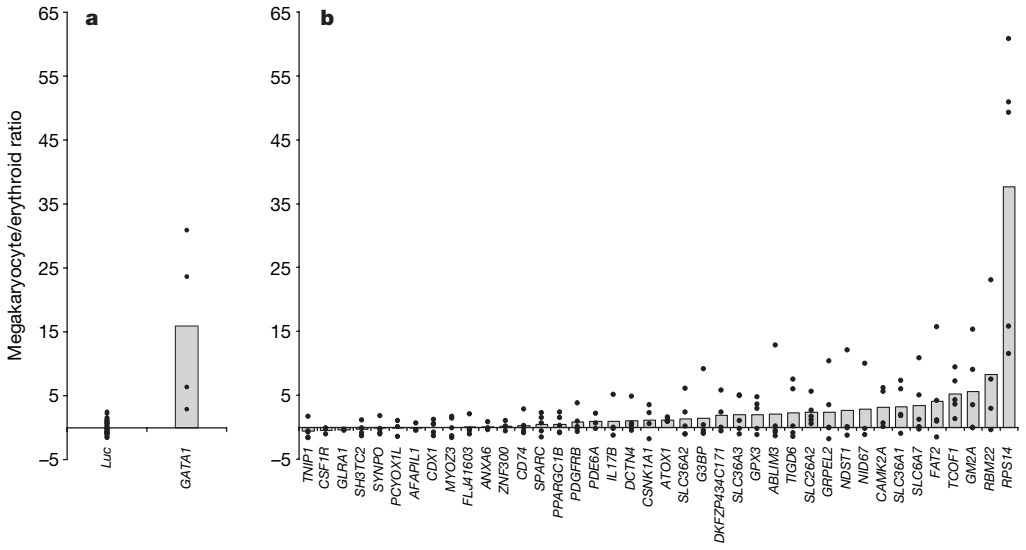

Figure 1 | Screen of the common deleted region for the 5q$^-$ syndrome. Each gene was targeted by multiple lentivirally expressed shRNAs in CD34$^+$ cells from umbilical cord blood, and the ratio of megakaryocytic to erythroid differentiation was determined by flow cytometry with antibodies against CD41 and GlyA, respectively. **a**, Controls: an shRNA targeting the luciferase gene (*Luc*), which is not expressed in the primary cells, and multiple shRNAs targeting *GATA1*, encoding an erythroid-specific transcription factor. **b**, All of the genes in the CDR for the 5q$^-$ syndrome. The megakaryocytic/erythroid ratio is shown as a *z*-score using the mean and standard deviation of the control (luciferase) replicates. For the control shRNA targeting the luciferase gene, circles represent 30 individual replicates. For all other genes, circles represent the median of three replicates for each individual shRNA. The mean of all shRNAs targeting a given gene is shown by the height of the grey bar.

macrocytosis, relative preservation of the platelet and neutrophil counts, erythroid hypoplasia in the bone marrow and an increased risk of leukaemia. Analogous to our results demonstrating RPS14 function in 18S pre-rRNA processing and 40S polysome formation, a similar requirement of RPS19 in ribosomal biogenesis has recently been shown[13]. Beyond Diamond–Blackfan anaemia, the genes implicated in other paediatric bone marrow failure syndromes, including Shwachman–Diamond syndrome, dyskeratosis congenita and cartilage–hair hypoplasia, are also involved in ribosomal biogenesis[14]. Our findings therefore establish a logical link between the 5q$^-$ syndrome, caused by the somatic deletion of one allele of *RPS14*, and congenital bone marrow failure syndromes, caused by the heritable mutation of other ribosome-associated proteins.

The erythroid specificity of acquired or inherited defects in RPS14, RPS19 or RPS24 expression is noteworthy. Although these ribosomal proteins and the ribosomal subunits they constitute are thought to be ubiquitous, the erythroid lineage is under particularly high biosynthetic demand. Erythroid progenitor cells proliferate extraordinarily rapidly (yielding 2×10^{11} new red cells per day in an adult human)[15], and contain extremely high concentrations of globin proteins—all resulting in a high demand for ribosomal biogenesis. Furthermore, erythroid cells must balance the production of haem and the translation of globin proteins precisely; otherwise the cells undergo apoptosis[16]. It is therefore possible that partial loss of ribosomal function in other lineages may not result in an obvious phenotype. We also note that in an unbiased screen in zebrafish for genes that cause tumours after the loss of a single allele, 92% of the tumour-prone fish lines had hemizygous mutations in genes encoding ribosomal proteins[17]. These observations suggest that loss of function of *RPS14* in the 5q$^-$ syndrome may explain not only the erythroid differentiation defect seen in affected patients but also their propensity to progress to acute leukaemia. The mechanism by which ribosomal dysfunction is tumorigenic in fish has yet to be determined.

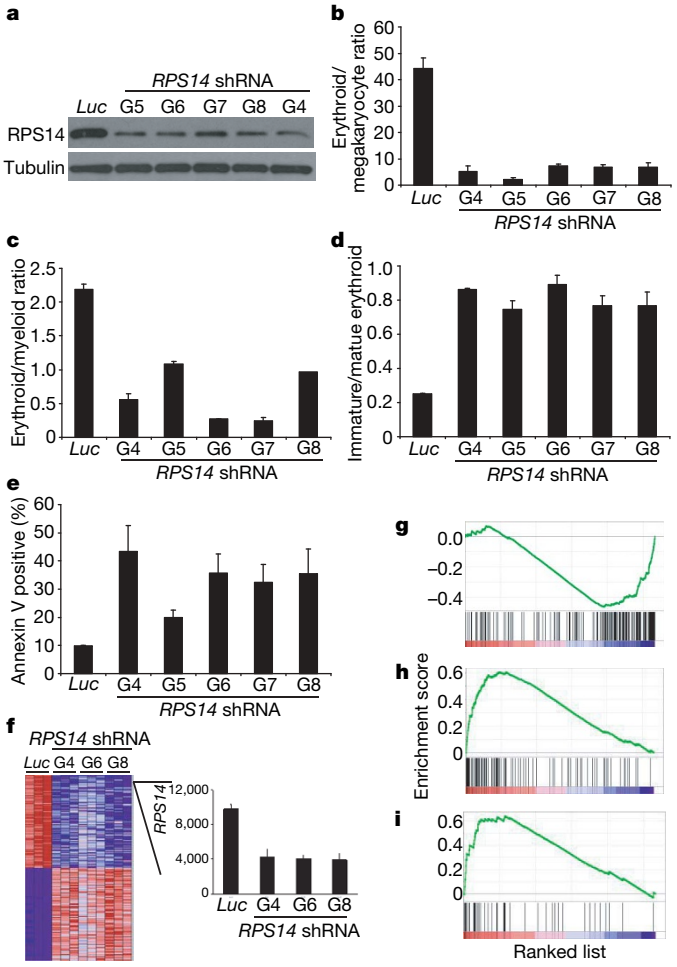

Figure 2 | Multiple shRNAs targeting *RPS14* recapitulate the 5q$^-$ syndrome *in vitro*. **a**, Western blots demonstrate that five different shRNAs effectively decrease levels of RPS14. **b**, In comparison with a control shRNA targeting the luciferase gene (*Luc*), each of the five *RPS14* shRNAs blocks erythroid relative to megakaryocytic differentiation in adult bone marrow CD34$^+$ cells. The ratios of cells from the erythroid and megakaryocytic lineages, indicated on the *y* axis, were assessed by flow cytometry with antibodies against GlyA and CD41, respectively. **c**–**e**, In addition, *RPS14* shRNAs decrease erythroid relative to myeloid differentiation, assessed with antibodies against GlyA and CD11b (**c**), block terminal erythroid differentiation, assessed with antibodies against GlyA and CD71 (**d**), and increase apoptosis, assessed by annexin V expression (**e**). In **b**–**e** the effect of *RPS14* shRNAs, in contrast with the *Luc* shRNA, was statistically significant ($P < 0.05$ by Student's two-tailed *t*-test, mean and s.e.m. shown ($n = 3$)). **f**, Multiple shRNAs targeting *RPS14* also alter the transcriptional programs of lineage-specific differentiation. The top 100 marker genes that are differentially expressed between cells expressing control versus *RPS14* shRNAs, ranked by signal to noise ratio[25]. *RPS14* is at the top of the list of downregulated genes and is expressed at about 40% of the normal level. Error bars indicate s.e.m. **g**–**i**, *RPS14* shRNAs significantly decrease the expression of an erythroid gene expression signature[26] (**g**) and increase the expression of neutrophil[27] (**h**) and platelet[28] (**i**) signatures, as assessed by GSEA. Genes are ranked by signal/noise ratio according to their differential expression between cells expressing *RPS14* and control shRNAs. Genes in the lineage-specific gene sets are marked with vertical bars, and the enrichment score is shown in green.

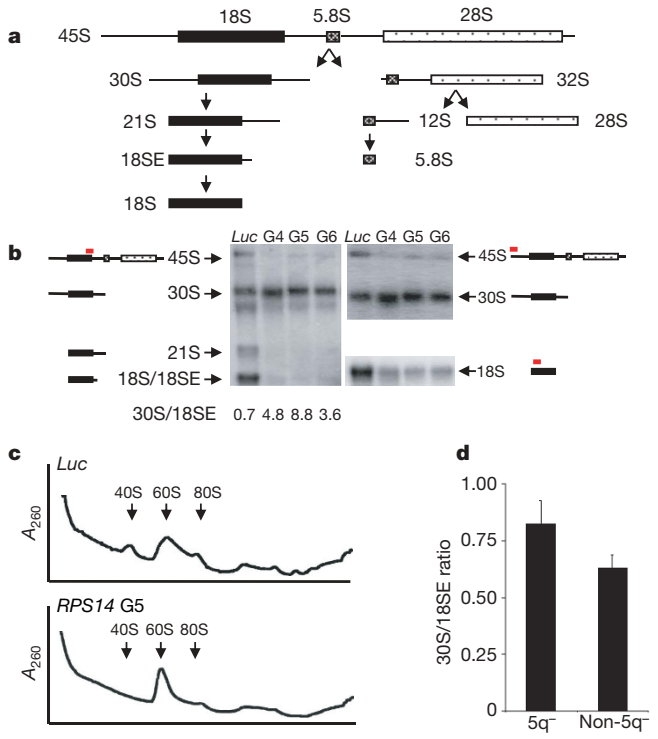

Figure 3 | RPS14 is required for 18S pre-rRNA processing and 40S ribosomal subunit formation. **a**, A simplified diagram of pre-rRNA processing. **b**, A defect in the 5′ processing of 18S pre-rRNA is evident from northern blots using RNA from TF-1 cells expressing control or *RPS14* shRNAs, with an accumulation of 30S rRNA and a deficiency of 21S and 18SE pre-rRNAs and mature 18S rRNA. The northern blot probes are shown in red. **c**, Polysome profiles from TF-1 cells show that decreased expression of RPS14 results in a 40S subunit deficiency. **d**, The 30S/18SE pre-rRNA ratio is also increased in RNA from bone marrow mononuclear cells from MDS patients with 5q$^-$ syndrome ($n = 4$) compared with that from MDS patients without 5q deletions ($n = 5$), as measured by quantification of northern blots ($P = 0.06$). Error bars indicate s.e.m.

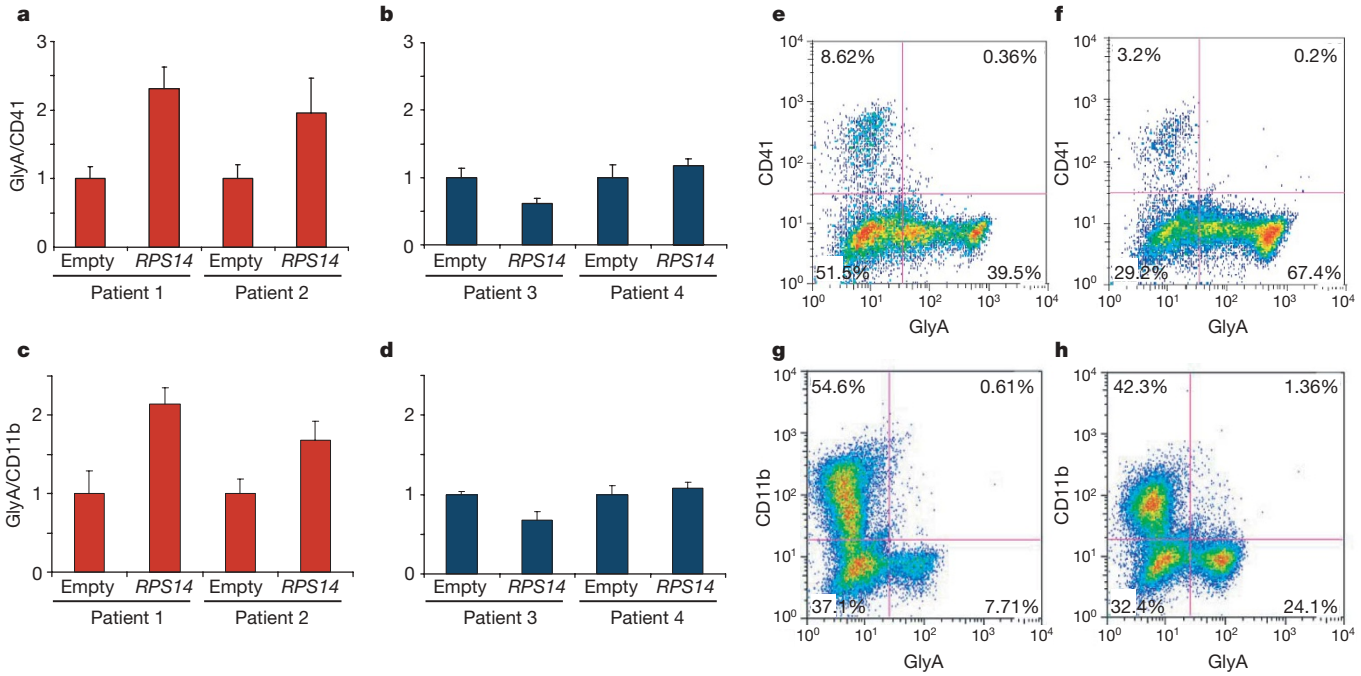

Figure 4 | RPS14 overexpression rescues erythroid differentiation in samples from patients with 5q deletions. **a–d**, $CD34^+$ cells from bone marrow aspirates of patients with the $5q^-$ syndrome (**a**, **c**; red) and MDS patients without 5q deletions (**b**, **d**; blue) were infected with a lentivirus expressing the *RPS14* complementary DNA or an empty vector. In patients with 5q deletions, RPS14 overexpression increased erythroid relative to megakaryocytic differentiation (**a**, **b**) and erythroid relative to myeloid differentiation (**c**, **d**) shown normalized to the empty-vector control. Means and s.e.m. for three independent experiments are shown. **e–h**, Representative flow cytometry plots for patient 1. In comparison with the empty vector control (**e**, **g**), overexpression of RPS14 (**f**, **h**) results in an increase in GlyA expression and a decrease in CD41 (**e**, **f**) and CD11b (**g**, **h**).

The experiments described here establish *RPS14* as a causal gene for the $5q^-$ syndrome. However, it is conceivable that other genes (on 5q or elsewhere) collaborate with *RPS14* to cause the disease phenotype. We speculate that whereas RPS14 loss of function may be sufficient for the erythroid differentiation defect, additional mutations may be required for RPS14-deficient cells to reach clonal dominance and to progress to malignant transformation to AML. In that regard, the $5q^-$ syndrome region on chromosome 5 should be distinguished from a more centromeric locus on 5q that has been associated with therapy-related and aggressive subtypes of MDS as well as AML, and for which two candidate genes have been recently reported[18–20]. In most patients a large portion of 5q is deleted, encompassing both critical regions, so it is possible that the loss of both *RPS14* and a second collaborating gene is achieved in a single genetic event.

Acquired deletions are a hallmark of cancer and pre-cancerous states. In general, such deletions flag the existence of a tumour suppressor gene conforming to Knudson's two-hit hypothesis, in which one allele is often deleted and the other allele is inactivated by deletion, mutation or epigenetic modification. However, in multiple tumour types (for example 1p deletions in neuroblastoma, 3p deletions in lung cancer, and 7q deletions in myeloid malignancies) the search for the key tumour suppressor gene has been elusive. A possible explanation for the failure to identify these classic tumour suppressor genes is that oncogenesis is caused by allelic insufficiency[21]. The recent discovery of monoallelic deletions or mutations in *PAX5* in acute lymphoblastic leukaemia supports this hypothesis[22]. Our RNA interference-based discovery of the $5q^-$ syndrome gene suggests that haploinsufficient disease genes can be identified with this approach. It is possible that the systematic application of RNAi might similarly identify the genes responsible for other diseases caused by allelic insufficiency.

METHODS SUMMARY

Culture of haematopoietic progenitor cells. Primary normal human bone marrow or umbilical cord blood $CD34^+$ cells were differentiated *in vitro* with a two-phase liquid culture system using combinations of cytokines supporting erythroid, myeloid and megakaryocytic differentiation[23]. Viable cells from bone marrow aspirates from patients with MDS were collected under a protocol approved by the institutional review board at Massachusetts General Hospital.

Lentiviral vectors. Multiple shRNA lentiviruses targeting each gene in the CDR for the $5q^-$ syndrome were produced as described previously[24]. The target sequence of each shRNA is listed in Supplementary Table 2.

Flow cytometry. Haematopoietic differentiation was assessed by flow cytometry with antibodies specific for terminally differentiated erythroid cells (GlyA), immature erythroid cells (CD71), megakaryocytes (CD41) and myeloid cells (CD11b).

Microarrays with GSEA. Linear amplification of RNA was performed with the Ovation kit (Nugen) and labelled cDNA was applied to oligonucleotide microarrays (Affymetrix). GSEA was performed as described previously[9]. Microarray experiments and gene sets are listed in Supplementary Tables 3 and 4, respectively, and the data are available at GEO under accession number GSE9487.

Ribosomal RNA processing and polysome profiles. The effect of *RPS14* knockdown on pre-rRNA processing was performed by northern blot analysis. Polysome fractionation on a sucrose gradient and spectrophotometric detection were performed as described previously[13].

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 24 August; accepted 16 November 2007.

1. Knudson, A. G. Jr. Mutation and cancer: statistical study of retinoblastoma. *Proc. Natl Acad. Sci. USA* **68**, 820–823 (1971).
2. Van den Berghe, H. *et al.* Distinct haematological disorder with deletion of long arm of no. 5 chromosome. *Nature* **251**, 437–438 (1974).
3. Heaney, M. L. & Golde, D. W. Myelodysplasia. *N. Engl. J. Med.* **340**, 1649–1660 (1999).
4. Giagounidis, A. A., Germing, U. & Aul, C. Biological and prognostic significance of chromosome 5q deletions in myeloid malignancies. *Clin. Cancer Res.* **12**, 5–10 (2006).
5. List, A. *et al.* Lenalidomide in the myelodysplastic syndrome with chromosome 5q deletion. *N. Engl. J. Med.* **355**, 1456–1465 (2006).
6. Boultwood, J. *et al.* Narrowing and genomic annotation of the commonly deleted region of the $5q^-$ syndrome. *Blood* **99**, 4638–4641 (2002).
7. Raza, A. *et al.* Apoptosis in bone marrow biopsy samples involving stromal and hematopoietic cells in 50 patients with myelodysplastic syndromes. *Blood* **86**, 268–276 (1995).

8. Zender, L. *et al.* Identification and validation of oncogenes in liver cancer using an integrative oncogenomic approach. *Cell* **125,** 1253–1267 (2006).
9. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102,** 15545–15550 (2005).
10. Ferreira-Cerca, S. *et al.* Roles of eukaryotic ribosomal proteins in maturation and transport of pre-18S rRNA and ribosome function. *Mol. Cell* **20,** 263–275 (2005).
11. Draptchinskaia, N. *et al.* The gene encoding ribosomal protein S19 is mutated in Diamond–Blackfan anaemia. *Nature Genet.* **21,** 169–175 (1999).
12. Gazda, H. T. *et al.* Ribosomal protein S24 gene is mutated in Diamond–Blackfan anemia. *Am. J. Hum. Genet.* **79,** 1110–1118 (2006).
13. Flygare, J. *et al.* Human RPS19, the gene mutated in Diamond–Blackfan anemia, encodes a ribosomal protein required for the maturation of 40S ribosomal subunits. *Blood* **109,** 980–986 (2007).
14. Liu, J. M. & Ellis, S. R. Ribosomes and marrow failure: coincidental association or molecular paradigm? *Blood* **107,** 4583–4588 (2006).
15. Quesenberry, P. J. & Colvin, G. A. in *Williams Hematology* 153 (McGraw-Hill, New York, 2005).
16. Quigley, J. G. *et al.* Identification of a human heme exporter that is essential for erythropoiesis. *Cell* **118,** 757–766 (2004).
17. Amsterdam, A. *et al.* Many ribosomal protein genes are cancer genes in zebrafish. *PLoS Biol.* **2,** E139 (2004).
18. Horrigan, S. K. *et al.* Delineation of a minimal interval and identification of 9 candidates for a tumor suppressor gene in malignant myeloid disorders on 5q31. *Blood* **95,** 2372–2377 (2000).
19. Liu, T. X. *et al.* Chromosome 5q deletion and epigenetic suppression of the gene encoding α-catenin (*CTNNA1*) in myeloid cell transformation. *Nature Med.* **13,** 78–83 (2007).
20. Joslin, J. M. *et al.* Haploinsufficiency of *EGR1*, a candidate gene in the del(5q), leads to the development of myeloid disorders. *Blood* **110,** 719–726 (2007).
21. Fodde, R. & Smits, R. Cancer biology. A matter of dosage. *Science* **298,** 761–763 (2002).
22. Mullighan, C. G. *et al.* Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* **446,** 758–764 (2007).
23. Ebert, B. L. *et al.* An RNA interference model of RPS19 deficiency in Diamond–Blackfan anemia recapitulates defective hematopoiesis and rescue by dexamethasone: identification of dexamethasone-responsive genes by microarray. *Blood* **105,** 4620–4626 (2005).
24. Moffat, J. *et al.* A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* **124,** 1283–1298 (2006).
25. Golub, T. R. *et al.* Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286,** 531–537 (1999).
26. Ebert, B. L. *et al.* An erythroid differentiation signature predicts response to lenalidomide in myelodysplastic syndrome. *PLoS Med.* (in the press).
27. Stegmaier, K. *et al.* Gene expression-based high-throughput screening (GE-HTS) and application to leukemia differentiation. *Nature Genet.* **36,** 257–263 (2004).
28. Gnatenko, D. V. *et al.* Transcript profiling of human platelets using microarray and serial analysis of gene expression. *Blood* **101,** 2285–2293 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank Broad Institute RNAi and Genetic analysis platforms for advice, single-nucleotide polymorphism analysis and reagents. This work was supported by grants from the National Heart Lung and Blood Institute to T.R.G., B.L.E. and S.R.E. T.R.G. is an investigator of the Howard Hughes Medical Institute.

Author Contributions B.L.E., J.P., J.B., C.Y.C., P.T. and S.R.E. performed experiments and analysed data. D.E.R. provided essential reagents. N.G., A.R. and E.A. provided samples from patients. B.L.E. and T.R.G. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to T.R.G. (golub@broad.mit.edu).

METHODS

Culture of haematopoietic progenitor cells. Cryopreserved human bone marrow CD34+ cells (Poietics) were obtained from Cambrex. Umbilical cord blood was harvested under a protocol approved by the institutional review board (IRB) at Brigham and Women's Hospital, and $CD34^+$ cells were purified using $CD34^+$ MACS microbeads (Miltenyi Biotec). Viable cells from bone marrow aspirates from MDS patients were banked under an IRB-approved protocol at Massachusetts General Hospital. To induce erythroid differentiation, cells were cultured in Serum-Free Expansion Medium (Stem Cell Technologies) supplemented with 100 U ml^{-1} penicillin/streptomycin, 2 mM glutamine, 40 μg ml^{-1} lipids (Sigma), 100 ng ml^{-1} stem cell factor, 10 ng ml^{-1} interleukin-3, 10 ng ml^{-1} interleukin-6 and 0.5 U ml^{-1} erythropoietin. The concentration of erythropoietin was increased to 3 U ml^{-1} on day 7. To support both erythroid and megakaryocytic differentiation in a single liquid culture, 50 ng ml^{-1} thrombopoietin was added to the culture. To support both erythroid and myeloid differentiation in a single liquid culture, 15 ng ml^{-1} granulocyte colony-stimulating factor (Neupogen; Amgen) and 40 ng ml^{-1} FLT-3 ligand were added. Cells were harvested for flow cytometry after 10 days of liquid culture.

Culture of TF-1 cells. TF-1 cells were maintained in RPMI medium supplemented with 10% fetal bovine serum, 100 U ml^{-1} penicillin/streptomycin, 2 mM glutamine and 1 ng ml^{-1} granulocyte–macrophage colony-stimulating factor. Doxorubicin and staurosporine were obtained from Calbiochem.

Lentiviral vectors and infection. Oligonucleotides encoding shRNAs were cloned into pLKO.1 as described previously[24]. Sequences targeted by each shRNA are listed in Supplementary Table 2. The *RPS14* cDNA was cloned into pLenti6.2/V5-DEST (Invitrogen). Lentiviral backbone vector and packaging plasmids were transfected into 293T cells, and viral supernatant was harvested as described previously[24]. Primary haematopoietic cells were infected with lentivirus one day after being thawed in the presence of 2 μg ml^{-1} Polybrene (Sigma) and selected 24 h later with 2 μg ml^{-1} puromycin (Sigma) for shRNA lentiviruses, and with 3 μg ml^{-1} blasticidin for cDNA-expressing lentiviruses.

Flow cytometry. Lineage-specific differentiation was evaluated by flow cytometry. About 5×10^5 cells were incubated for 15 min on ice with phycoerythrin, phycoerythrin-Cy5 or fluorescein isothiocyanate-conjugated antibodies against glycophorin-A (CD235a, clone GA-R2; BD Pharmingen), CD71 (clone M-A712; BD Pharmingen), CD11b (clone ICRF44; BD Pharmingen), CD41 (clone HIP8; BD Pharmingen) or annexin V.

Gene expression profiling. RNA was purified from mononuclear cells with the use of Trizol (Invitrogen). Linear amplification of 20 ng of total RNA was performed with the Ovation Biotin RNA Amplification and Labelling System (Nugen). Fragmented, labelled cDNA was hybridized to HG_U133AAofAv2 microarrays (Affymetrix). Raw expression values were normalized by using robust multiarray averaging[29]. Marker genes were ranked with the signal/noise metric[25]; For gene x this metric, S_x, is calculated as

$$S_x = (\mu_0 - \mu_1)/(\sigma_0 + \sigma_1)$$

where μ_0 and σ_0 are the mean and standard deviation for gene x in class 0, and μ_1 and σ_1 are the respective values for class 1. All microarray experiments are listed in Supplementary Table 3. The complete data set, along with Supplementary Information, is available at http://www.broad.mit.edu/cancer/pub/5qMDS.

GSEA was performed as described previously[9]. Erythroid-specific genes were defined by genes that are increased during terminal erythroid differentiation *in vitro*[26]; neutrophil-specific genes were defined by comparing mature neutrophils with primary AML blast cells[27]; a platelet-specific gene set was defined previously[28]; and a lenalidomide signature, developed previously, was defined by the genes that are expressed at significantly higher levels in bone marrow mononuclear cells from patients who do not respond to lenalidomide, compared with patients who do respond to the drug[26]. Statistical significance was assessed by random permutation of the gene sets[9]. All gene sets are listed in Supplementary Table 4.

Western blots. Western blots were performed as described previously, using antibodies against RPS14 (A01; Abnova) at 1:500 dilution and antibodies against α-tubulin (Ab-2; Neomarkers) at 1:1,000 dilution. Image analysis was performed with ImageJ software (http://rsb.info.nih.gov/nih-image).

Ribosomal RNA analysis. Total RNA was isolated from TF-1 cells or patient samples by using Trizol (Invitrogen). RNA was fractionated on 1.5% formaldehyde–agarose gels and transferred to Zetaprobe membrane (Bio-Rad). Membranes were washed overnight at 55 °C with 2 × SSC (0.3 M NaCl, 0.03 M sodium citrate, pH 7.0) and 1% SDS and prehybridized for a minimum of 4 h with ULTRAhyb oligonucleotide hybridization buffer (Ambion). The oligonucleotide probes used were as follows: 5′ ETS, 5′-ACCGGTCACGACTCGGCA-3′ (complementary to sequences 1786–1804 in 5′ ETS of the ribosomal RNA transcription unit); 18S, 5′- GCATGGCTTAATCTTTGAGACAAGCATAT-3′ (complementary to sequences 3681–2709 in 18S rRNA); and 18S/ITS1, 5′-CCTCGCCCTCCGGGCTCCGTTAATGATC-3′ (complementary to sequences 5520–5547 spanning the boundary between 18S rRNA and ITS1). The oligonucleotides, at a concentration of 30 pM, were labelled with [γ-^{32}P]ATP by using T4 polynucleotide kinase (New England Biolabs). Membranes were hybridized overnight at 37 °C in ULTRAhyb oligonucleotide hybridization buffer and washed the following morning three times with 6 × SSC at 37 °C. Washed membranes were subjected to phosphorimage analysis (Phosphorimager SF; Molecular Dynamics) for quantification.

Polysome analysis. Extracts from TF-1 cells infected with *RPS14* or control shRNAs were prepared as described previously[30]. Extracts were layered on 16-ml 15–55% sucrose gradients and centrifuged in a SW28.1 rotor (Beckman Instruments) for 5 h at 28,000 r.p.m. Gradients were fractionated, and A_{254} was monitored on an ISCO model 185 gradient fractionator using a UA-6 absorbance detector.

Statistical analysis. The significance of experimental results was determined by Student's *t*-test unless otherwise noted. The significance of RPS14 overexpression relative to control, in samples from patients with 5q deletions compared with patients without 5q deletions, was determined by a two-way analysis of variance.

For the shRNA screen of genes in the 5q CDR, the likelihood that each gene significantly altered differentiation was determined by using a modified Kolmogorov–Smirnov statistic, similarly to the procedure implemented in GSEA[9]. For each gene, the set of shRNAs targeting that gene were combined into a gene set. All scores for the screen were sorted to create a ranked list. The enrichment score of each gene set was calculated by using a modified Kolmogorov–Smirnov statistic. In brief, the enrichment score is computed as a Kolmogorov–Smirnov statistic, namely, the maximum deviation from zero of the difference between the empirical cumulative distribution function (ECDF) of probe scores for a given gene, and the ECDF of the probe scores of all the other genes. Bonferroni *P* values were calculated to correct for multiple hypotheses.

29. Bolstad, B. M., Irizarry, R. A., Astrand, M. & Speed, T. P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185–193 (2003).
30. Tang, H. *et al.* Amino acid-induced translation of TOP mRNAs is fully dependent on phosphatidylinositol 3-kinase-mediated signaling, is partially inhibited by rapamycin, and is independent of S6K1 and rpS6 phosphorylation. *Mol. Cell. Biol.* **21**, 8671–8683 (2001).