HHS Public Access

Author manuscript

Nat Genet. Author manuscript; available in PMC 2012 April 01.

Published in final edited form as:

Nat Genet. ; 43(10): 1012–1017. doi:10.1038/ng.913.

Heritable *GATA2* Mutations Associated with Familial Myelodysplastic Syndrome and Acute Myeloid Leukemia

Christopher N. Hahn[1,2], **Chan-Eng Chong**[1,2,15], **Catherine L. Carmichael**[3,15], **Ella J. Wilkins**[3,13], **Peter J. Brautigan**[1], **Xiao-Chun Li**[1], **Milena Babic**[1], **Ming Lin**[1], **Amandine Carmagnac**[3], **Young K. Lee**[1], **Chung H. Kok**[4,5], **Lucia Gagliardi**[1], **Kathryn L. Friend**[6], **Paul G. Ekert**[7], **Carolyn M. Butcher**[4,5], **Anna L. Brown**[5], **Ian D. Lewis**[2,5], **L. Bik To**[2,5], **Andrew E. Timms**[8], **Jan Storek**[9], **Sarah Moore**[1], **Meryl Altree**[10], **Robert Escher**[3,14], **Peter G. Bardy**[5], **Graeme K. Suthers**[10,11], **Richard J. D'Andrea**[2,4,5,16], **Marshall S. Horwitz**[8], and **Hamish S. Scott**[1,2,3,12,16]

[1]Department of Molecular Pathology, Centre for Cancer Biology, SA Pathology, Adelaide, SA, Australia

[2]School of Medicine, University of Adelaide, SA, Australia

[3]Molecular Medicine Division, Walter and Eliza Hall Institute of Medical Research, Parkville, Vic, Australia

[4]Department of Haematology and Oncology, The Queen Elizabeth Hospital, Woodville, SA, Australia

[5]Department of Haematology, Centre for Cancer Biology, SA Pathology, Adelaide, SA, Australia

[6]Department of Paediatric and Reproductive Genetics, SA Pathology, Adelaide, SA, Australia

Correspondence should be addressed to H.S.S. (hamish.scott@health.sa.gov.au).

[13]Present Address: Neurogenetics Laboratory, Howard Florey Institute, Parkville, VIC, Australia.

[14]Present Address: Medical Clinic, Regional Hospital Emmental, Burgdorf, Switzerland.

[15]These authors contributed equally to this work.

[16]These authors jointly supervised this work.

URLs. dbSNP132, http://www.ncbi.nlm.nih.gov/SNP/; 1000 Genomes Project, http://www.1000genomes.org/; PolyPhen-2, http://genetics.bwh.harvard.edu/pph2/.

Accession numbers. Affymetrix expression data are available from Gene Expression Omnibus under accession GSE29276. Human *GATA2* cDNA sequence is available from Genbank under accession number NM_032638.4.

AUTHOR CONTRIBUTIONS

C.N.H., R.J.D., M.S.H. and H.S.S. managed the project. C.N.H., C.L.C., E.J.W., C-E.C., P.J.B., X-C.L., M.S., M.L., A.C., Y.K.L., C.M.B., K.L.F. and A.E.T. performed the experiments. C.H.K. and R.J.D. performed structural modeling and C.N.H., C.H.K. and L.G. performed data analysis. L.B.T., M.A., J.S., P.G.B., G.K.S., R.J.D., M.S.H. and H.S.S. collected families with MDS/AML, and provided clinical data and samples. R.E. and P.G.E. participated in experimental design and provided critical reagents. A.L.B., I.D.L., S.M., L.B.T. provided sporadic AML samples and correlative clinical data. C.N.H., R.J.D., M.S.H. and H.S.S. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

While this manuscript was under consideration, Hsu et al have identified germline *GATA2* mutations, including p.Thr354Met and other ZF2 mutations in "MonoMAC syndrome", an autosomal dominant syndrome associated with myelodysplasia and myeloid leukemias as well as monocytopenia, B and NK cell lymphopenia and mycobacterial, fungal and viral infections[28]. We also have reason to believe that the syndromes described by Mansour et al[29] and Bigley et al[30] are also related to germline mutations in *GATA2*.

[7]Cell Signalling and Cell Death Division, Walter and Eliza Hall Institute of Medical Research, Parkville, Vic, Australia

[8]Department of Pathology, University of Washington School of Medicine, Seattle, WA, USA

[9]Department of Medicine, University of Calgary, Calgary, Alberta, Canada

[10]SA Clinical Genetics Service, SA Pathology, Adelaide, SA, Australia

[11]Department of Paediatrics, University of Adelaide, Adelaide, SA, Australia

[12]School of Molecular and Biomedical Science, University of Adelaide, SA, Australia

Abstract

We report the discovery of the *GATA2* gene as a new myelodysplastic syndrome (MDS)/acute myeloid leukemia (AML) predisposition gene. We found the same, novel heterozygous c.1061C>T (p.Thr354Met) missense mutation in the *GATA2* transcription factor gene segregating with the multigenerational transmission of MDS/AML in three families, and a *GATA2* c.1063_1065delACA (p.Thr355del) mutation at an adjacent codon in a fourth MDS/AML family. The mutations reside within the second zinc finger of GATA2 which mediates DNA-binding and protein-protein interactions. We show differential effects of the mutants on transactivation of target genes, cellular differentiation, apoptosis and global gene expression. Identification of such predisposing genes to familial forms of MDS and AML is critical for more effective diagnosis and prognosis, counselling, selection of related bone marrow transplant donors, and development of therapies.

AML is the most common form of sporadic leukemia in adults [1] while MDS is a clonal disorder of hematopoietic stem cells characterized by ineffective hematopoiesis, with a tendency to progress to AML[2]. The study of families predisposed to particular malignancies is a successful strategy for discovering causative oncogenes and tumour suppressor genes (TSG). While rare, dozens of families developing non-syndromic forms of MDS and AML (*i.e.* lacking other systemic manifestations) have been described. To date, only two MDS/AML predisposition genes have been recognized: Runt-related transcription factor 1 (*RUNX1*/*AML1*) and CCAAT-enhancer binding protein α (*CEBPA*) (reviewed in[3]).

Here we report a highly specific p.Thr354Met heritable mutation in GATA2 co-segregating with with early onset MDS/AML in three families. We also report a family with MDS with a 3 bp heritable deletion in the *GATA2* gene (p.Thr355del) deleting the second threonine in this sequence.

We determined the genomic DNA sequence of all RefSeq exons in 50 candidate genes (Supplementary Table 1) from patients representing five pedigrees with predisposition to MDS/AML, prescreened for absence of *RUNX1* or *CEBPA* germline mutations. In three families, there was an identical heritable heterozygous variation in the transcription factor *GATA2*, c.1061C>T (p.Thr354Met) (Fig. 1, Supplementary Fig. 1). In all three families p.Thr354Met segregated with the disease for the samples tested, and no family members had AML or MDS who did not also carry p.Thr354Met (Fig. 1a, Supplementary table 2). There

were also members in each family who carried this variant but were unaffected (Pedigree 1: III-5 and III-8; Pedigree 2: II-6; Pedigree 3: III-9).

We recently identified a fourth family in which a father and son, both affected by MDS, shared a heterozygous heritable deletion of 3 bp in *GATA2* (c.1063-1065delACA) resulting in p.Thr355del (Fig. 1, Supplementary Fig. 1). This codon is adjacent to the codon mutated in the first three families, and encodes the second of the five consecutive threonines.

p.Thr354 and p.Thr355del are the first two of five consecutive threonine residues in a highly conserved region of the GATA2 protein (Supplementary Fig. 1b) encoding zinc finger 2 (ZF2), which is involved in DNA binding, homodimerization and interaction with transcription factor PU.1[4,5]. PolyPhen-2 predicts that p.Thr354Met and p.Thr355del are likely to affect GATA2 function. Somatic mutations in ZF2 of GATA2 have also been reported during chronic myeloid leukaemia (CML) blast crisis (BC) (p.Leu359Val, p.Ala341_Gly346del)[6] and recently in ZF1 and ZF2 in AML-M5[7] (Fig. 1b,c). Somatic mutations in the corresponding ZF2 of the related protein family member GATA3 are found in breast cancer[8] (Fig. 1c).

High resolution melt (HRM) analysis did not detect p.Thr354Met, p.Thr355del or other variants in exon 5 of *GATA2* in 695 non-leukemic, ethnically-matched normal controls (*i.e.* 1390 chromosomes) (Supplementary Fig. 2a, Supplementary Note). Thus it is improbable that these variants represent rare polymorphisms. These variants were also not present in dbSNP132 or 1000 Genomes Project (January, 2011; URLs). Together with the disease segregation data, these results indicate that the GATA2 p.Thr354Met and p.Thr355del variants are the predisposing mutation in these families with familial MDS/AML.

A distinguishing feature of our families with *GATA2* mutation was a lack of apparent “accessory” phenotype inside or outside the hematopoietic system, akin to the thrombocytopenia and eosinophilia seen in AML-predisposed families due to *RUNX1* and *CEBPA* mutations, respectively[9,10]. In all 4 families, the *GATA2* mutations were associated with early-onset MDS and/or AML displaying highly penetrant autosomal dominant inheritance and resulted in a poor outcome unless successfully transplanted (*e.g.* Pedigree 1: age of death from AML – 10 – 50 years; 2 individuals of age 58 and 62 years have mutation but no disease; all other siblings without the mutation are alive or have lived beyond 53 years) (Supplementary Table 2 and 3, Supplementary Note). For Pedigrees 1 and 3[11,12], the presentation varied with some displaying protracted MDS and others acute onset; the FAB subtype and karyotypic features of AML varied. In Pedigree 4, MDS was first diagnosed at age 13 in the son, who was treated with allogeneic bone marrow transplant at age 15; MDS was later diagnosed at age 53 in the father, who underwent allogeneic bone marrow transplantation.

Heritable *GATA2* coding variations were not found in samples from another 8 families with multiple cases of AML, or in another 27 families with multiple occurrences of various lymphoid malignancies (11 NHL, 5 HL, 3 ALL, 7 CLL, 1 Multiple myeloma families). Also, no mutations were detected in *GATA2* in 15 hematopoietic cell lines (Supplementary Table 4). No sequence variations were detected in the entire *GATA2* coding region of 268

sporadic AML patient sample DNAs except a single c.182C>T (p.Ala61Val) variant in exon 2 (Supplementary Fig. 2c), which was assessed to be benign using PolyPhen-2 (URLs). Together, this suggests that point mutations and small indels in the *GATA2* coding sequence are not frequent in sporadic AML.

Haplotype mapping using 8 informative single nucleotide polymorphisms (SNPs) within and surrounding the *GATA2* gene demonstrated that the c.1061C>T (p.Thr354Met) mutation segregated within two distinct haplotypes (Supplementary Table 5) indicating that this mutation has arisen at least twice among the three families in which it is found.

GATA2 is a DNA-binding transcription factor which localizes predominantly to the nucleus. We generated cDNAs for the p.Thr354Met and p.Thr355del mutant GATA2 proteins and the acquired CML BC p.Leu359Val mutant[6]. Wildtype (WT) and mutant proteins expressed at comparable levels when transiently expressed in HEK293 fibroblasts (Fig. 2a) and when induced to express from a 4HT-responsive dual vector lentivirus system in stably transduced HL-60 promyelocytes (Supplementary Fig. 3a). Mutant proteins were expressed at comparable levels to wildtype (WT) GATA2 when transiently introduced into WT and mutant proteins appropriately localized to the nucleus (Fig. 2b, Supplementary Fig. 3b). However, the p.Thr354Met mutation dramatically reduced the ability of GATA2 to bind its consensus WGATAR DNA motif while p.Thr355del almost completely ablated DNA binding (Fig. 2c, Supplementary Fig. 3c).

Molecular modeling of GATA2 ZF2 (Supplementary Fig. 4, Supplementary Note) demonstrated that the p.Thr354 residue does not contact DNA, but rather makes polar contact with adjacent threonines, and via its amino group with p.Cys349 which coordinates the zinc atom. Replacement of p.Thr354 with the bulky methionine moiety is predicted to alter the overall structure of this zinc finger by affecting zinc contacts. This may explain reduced binding of p.Thr354Met to DNA (Fig. 2c, Supplementary Fig. 3c). In contrast, p.Thr355del shortens the conserved threonine string, likely impacting the orientation and position of p.Leu359, which directly contacts DNA. These observations likely explain the almost-complete ablation of DNA binding.

Luciferase reporter assay experiments show that GATA2 p.Thr354Met and p.Thr355del had significantly reduced transactivation ability compared to WT on known *GATA2* responsive enhancers (*RUNX1* and *CD34*) and the *LYL1* promoter (Fig. 3a,b,c). Experiments mixing WT with p.Thr354Met or p.Thr355del at a 1:1 ratio, mimicking heterozygosity, demonstrated a dominant negative effect of the mutants over WT transcription activation in multiple systems (Fig. 3d,e, Supplementary Fig. 5). Interestingly, WT and PU.1 transactivated the *CSF1R* (*M-CSF-R*) promoter 2.4 and 2.5-fold, respectively, but together synergized to induce 18-fold (Fig. 3e). While p.Leu359Val was similar to WT GATA2, p.Thr354Met and p.Thr355del gave dramatically reduced induction alone (1.5- and 0.9-fold) or with PU.1 (7- and 9-fold, respectively, compared to 18-fold with WT). p.Thr354Met and p.Thr355del also displayed dominant negative activity with transactivation by WT GATA2 reduced to 9- and 10-fold, respectively in the presence of these mutants (only marginally above the 7- and 9-fold with mutants alone). Hence, p.Thr354Met and p.Thr355del perturb the transactivation ability of GATA2, presumably by disrupting association with PU.1 or

other interacting transcription factors, and are likely impact expression of downstream targets. Interestingly, while WT GATA2 displayed different responses in HEK293 versus Cos-7 cells on the *RUNX1* enhancer (activating versus repressing, respectively), p.Thr354Met displayed loss-of-function activity in both cell types (Supplementary Fig. 6). Thus, on multiple GATA responsive elements, p.Thr354Met and p.Thr355del show loss-of-function and also dominant negative effects.

HL-60 promyelocytes differentiate into granulocytes upon exposure to all-*trans* retinoic acid (ATRA), resulting in upregulation of CD11b, cessation of proliferation and subsequent promotion of apoptosis (Fig. 4). When expressed at equivalent levels under non-differentiating conditions, unlike WT and p.Leu359Val which inhibited proliferation and promoted apoptosis, p.Thr354Met and p.Thr355del acted as loss-of-function mutants (Fig. 4b,f,j). However, in the presence of ATRA, p.Thr354Met alone enabled cell proliferation/survival (Fig. 4h), while simultaneously inhibiting differentiation and apoptosis (Fig. 4d,m, Supplementary Fig. 7). p.Thr355del appeared to be a null mutant under these conditions.

In order to better understand the effects of the GATA2 mutants on gene expression, microarray analysis was performed to compare global gene expression in HL-60 cells expressing WT GATA2 and the three GATA2 mutants (Supplementary Table 6 and Supplementary Fig. 8). The data clearly showed that p.Thr355del and p.Thr354Met are almost total loss-of-function mutants (Supplementary Fig. 8). Note that, p.Leu359Val exhibits gain-of-function (1,253 newly regulated genes compared to WT GATA2) and partial loss-of-function (457 genes no longer regulated) while retaining 786 genes commonly regulated. These results are consistent with EMSA-western blot and transactivation assays. Further bioinformatics analysis indicated that *MYC* may be among key target which is repressed by GATA2 WT but not p.Thr354Met and p.Thr355del (Supplementary Table 6, Supplementary Note)

Interestingly, recurrent p.Leu359Val mutation in ZF2 of GATA2 was reported in 8/85 cases of CML BC[6], a disease often phenotypically indistinguishable from AML. As shown in Fig. 1b,c, p.Thr354Met is situated between the deleted residues (p.Ala341_Gly346del), also observed in CML BC, and the p.Leu359 residue. p.Leu359 contacts DNA at the guanine residue of the WGATAR consensus motif and based on *in vitro* DNA binding and transactivation assays, p.Leu359Val has previously been reported to be a gain-of-function mutation while p.Ala341_Gly346del appears to be a partial loss-of-function mutation. p.Met354 or deletion of p.Thr355 may affect overall ZF structure (Supplementary Fig. 4) although we cannot exclude disrupted heterodimerization with GATA2's interacting partners of GATA2 (Supplementary Table 7). We speculate that aberrant protein partnerships may explain dominant negative activity and adversely influences expression of genes critical to myelopoiesis.

The MDS/AML observed within these families is clinically heterogeneous, and demonstrate a variety of somatic chromosomal abnormalities, including monosomy 7, trisomy 8, and trisomy 21 (Supplementary Table 2). As such it is similar to familial MDS/AML with monosomy 7[3,13]. Moreover, *GATA2* mutations were not detected in 8 MDS/AML families in which *RUNX1* and *CEPBA* mutations were excluded and 27 families presenting with

lymphoid malignancy. Mutations at this position within ZF2 are likely to initiate an exclusively myeloid pathway of oncogenesis in which subsequent gene-specific, somatically acquired mutations probably define the particular type of disease that ultimately arises.

The mechanism by which GATA2 p.Thr354Met and p.Thr355del mutations function is distinct to that generally described for *RUNX1* and *CEBPA*, which commonly act as classical TSG with a wide-range of mutations and requiring functional disruption of both alleles. Transcription factors are well characterized as targets of dominant negative or constitutively active mutations in cancer[14,15], with *RUNX1* mutations leading to a spectrum of outcomes including AML and ALL consistent with both TSG and dominant oncogene models[16,17]. While we have only been able to detect single allele *GATA2* germline mutations in affected samples, we cannot rule out the possibility of acquired mutations in the "normal" allele.

GATA2 is indispensable for hematopoiesis[17, 18,19]. It associates with, regulates or is regulated by transcription factors implicated in myeloid malignancy (Supplementary Table 7). Many of these interactions involve ZF2 in which the p.Thr354Met and p.Thr355del mutations reside, and it is likely that changes in the nature of these interactions play an important role in predisposition to MDS/AML. Indeed, our co-transfection studies are consistent with altered transactivation by p.Thr354Met and p.Thr355del with PU.1 (Fig. 3e).

p.Thr354Met, p.Thr355del or any other mutations in *GATA2* were absent in our heterogeneous cohort of sporadic AML patients, although we cannot rule out possible mutations in samples with low percentage blasts. This is consistent with other recent studies, however, suggesting that somatic *GATA2* mutations in both ZF1 and ZF2 could be acquired only in specific AML subtypes such as AML-M5[7,20,21]. *GATA2* is, however, overexpressed in many cases of sporadic MDS[22] and AML, particularly in *FLT3-ITD*$^{+}$ AML[23], suggesting that alterations to *GATA2* expression, rather than direct mutation, may occur more commonly.. Further, chromosomal aberrations at the 3q21 breakpoint cluster encompassing a presumptive *GATA2* regulatory region resulted in upregulated *GATA2* expression in MDS and AML[22,24–26]. In addition, retroviral insertional mutagenesis in *NUP98-HOXD13* mice, a model for MDS/AML, identified *Gata2* as a common insertion site in induced AML, all of which overexpressed *Gata2*[27]. Hence, accumulating evidence suggests that aberrant activation or overexpression of *GATA2* contributes to AML.

In this study, we show that *GATA2* is a new predisposition gene for familial MDS/AML and demonstrate functional changes due to mutations within a highly conserved threonine repeat located in the second zinc finger. Our findings highlight the power of approaches investigating familial predispositions to cancer, and have implications for diagnostic genetic testing. The poor outcome associated with these mutations may suggest that an aggressive treatment strategy is appropriate for individuals carrying *GATA2* mutations.

ONLINE METHODS

Patients

Families (Supplementary Table 2, Supplementary Note) were recruited and sample use approved through institutional human ethics review board approved protocols from the Australian Familial Haematological Cancer Study (Royal Adelaide Hospital (RAH) #091203 and #100702, and Children, Youth and Women's Health Service #REC1542/12/12, Adelaide, SA Australia), The Queen Elizabeth Hospital and the University of Washington (Seattle, WA USA).

Sequence analysis of candidate genes

To identify germline and somatic mutations in patients with familial AML, a panel of 50 hematopoietic candidate genes, incorporating a total of 638 exons, was assembled (Supplementary Table 1). Primer design, PCR amplification, and dideoxy sequencing of genomic DNA purified from lymphoblastoid cells of probands from 7 MDS/AML pedigrees were performed by the Australian Genome Research Facility (AGRF). Sequences were aligned with NCBI RefSeq sequences using Mutation Surveyor (SoftGenetics) and variants compared to the UCSC and NCBI SNP databases for novelty. Sequence changes were confirmed by re-sequencing in both directions. Primer sequences are available upon request. Screening of control and sporadic AML populations was performed using high resolution melt (HRM) analysis (Supplementary Table 8 and Supplementary Note).

Cell culture

HEK293, 293T and Cos-7 cells were cultured in DMEM with 10% fetal bovine serum (FBS) (JRH Biosciences) and transient transfections were performed using Lipofectamine 2000 (Invitrogen). HL-60 promyelocytic cells were cultured in RPMI containing 10% FBS. All cultures contained 50 units/ml penicillin and 50 μg/ml streptomycin (Sigma).

Generation of mutant *GATA2* plasmid and lentiviral expression constructs

An expression clone (pCMV6-XL6-GATA2) containing a 3.7 kb *GATA2* cDNA insert was obtained from OriGene, and p.Thr354Met, p.Thr355del and p.Leu359Val mutants were generated by site directed mutagenesis. The coding regions of wildtype (WT), p.Thr354Met and p.Leu359Val were cloned into a dual lentiviral vector system which was used to generate HL-60 cells expressing *GATA2* WT or mutants upon addition of 4-hydroxytamoxifen (4HT) (see Supplementary Note).

GATA2-responsive promoter and enhancer studies

The *GATA2*-responsive promoter (*LYL1*) and enhancer (*RUNX*) were PCR amplified and cloned into pGL4.12[*luc2CP*] (*Sfi*I) and pGL3-Promoter (*Kpn*I/*Bgl*II) (Promega), respectively. The CSF1R (M-CSF-R) promoter was PCR amplified and cloned into pGL4.12[*luc2CP*] (*Sfi*I). See Supplementary Table 9 for PCR primers used. The GATA2-responsive *CD34* enhancer-luciferase construct (CD34x2/Luc) and one with the GATA binding sites mutated (mutant CD34x2/Luc)[31] were kindly provided by Tariq Enver, Weatherall Institute of Molecular Medicine, Oxford, U.K. HEK293 or Cos-7 cells were

transfected at 90% confluence with Lipofectamine 2000. In all experiments, the molar equivalents of EV constructs were used to balance gene expressing constructs to avoid squelching artifacts. After 20 h, cells were harvested and luciferase activity determined with the Dual-Luciferase Reporter Assay System (Promega) using a GloMax®-Multi Detection System (Promega). All assays were performed a minimum of three times in triplicate. All results were analysed using Student's t-test, and reported as mean ± s.e.m. with significance, $p<0.05$ (asterisk).

Cell differentiation and proliferation assays

HL-60 cells were plated at 1.25×10^4 cells/ml and treated with or without 30 nm 4-hyrdoxy tamoxifen (4-HT) for 24 h and then with or without 2 μM all-*trans* retinoic acid (ATRA). Cell numbers were determined by manual counting and FACS analysis (Phycoerythrin anti-mouse CD11b and Phycoerythrin rat IgG2b isotype control) (eBioscience) was performed 6 days after addition of ATRA (Sigma). The cells were also stained with hematoxylin and eosin for assessing morphological changes.

Haplotyping

Haplotype mapping was performed by PCR amplification and sequencing of amplicons containing 50 single nucleotide polymorphisms (SNP) within and surrounding the position of the p.Thr354Met variant of the *GATA2* gene (Supplementary Table 10). All amplicons were generated using AmpliTaq Gold (Applied Biosystems) according to the manufacturer's protocol using 2 mM $MgCl_2$ and the following cycle strategy; 95°C, 10 min; 95°C, 30 s, 66°C – 58°C, 20 s (touchdown, 0.8°C/cycle for 10 cycles), 72°C, 45 s (total of 40 cycles); 72°C, 3 min.

Generation of mutant GATA2 plasmids and lentiviral expression constructs

An expression clone (pCMV6-XL6-GATA2) containing a 3.7 kb *GATA2* cDNA insert was obtained from OriGene (Cat. No. SC125368). p.Thr354Met, p.Thr355del and p.Leu359Val mutants were generated by QuikChange mutagenesis (Stratagene) using the primers T354M-F and T354M-R, 355delT-F and 355delT-R, and L359V-F and L359V-R (Supplementary Table 11), respectively. For the generation of lentiviral expression constructs, the regulatable pF 5xUAS W SV40 Puro (5xUAS)[32] was used. GATA2 WT or mutants were PCR amplified from the above pCMV6 plasmid vectors using the primers (KOZAK-GATA2-F and either GATA2-FLAG-R or GATA2-R) (Supplementary Table 12) and Pfu Turbo (Stratagene), excised with *Xba*I and cloned into the unique *Xba*I site of 5xUAS.

Generation of regulatable GATA2 expressing HL-60 cell lines

A dual lentiviral vector system was used to generate HL-60 cells expressing GATA2 WT or mutants upon addition of 4-hydroxytamoxifen (4HT) (Sigma). Infectious third generation lentivirus was made by cotransfecting 293T cells with either 5xUAS-GATA2 (WT or mutants) or pF GEV16 Super PGKHygro (GEV16)[33] plasmid and the three packaging plasmids pHCMVwhvgagpolml, pHCMV-G and pHCMVwhvrevml[34] (mass ratio 50:5:2.5:1). Supernatants were harvested 24 h later and filtered (Nalgene 45 μm syringe filter) (Nalge Nunc Int.). HL-60 cells were firstly transduced with GEV16 lentiviral

supernatant including 4 μg/ml polybrene and 2.5 μg/ml fungizone. After 48 h, HL-60GEV cells were selected in 1 mg/ml hygromycin (Roche). These cells were subsequently transduced with the GATA2 (WT, p.Thr354Met, p.Thr355del and p.Leu359Val) or EV lentiviral supernatant and selected in 3 μM Puromycin (Sigma).

Immunofluorescence staining

HL-60 cells carrying stably transduced 4HT-regulatable GATA2 (WT, p.Thr354Met, p.Thr355del and p.Leu359Val) were treated with and without 100nM 4HT. After 24 h, the cells were fixed with 4% of paraformaldehyde for 10 min. The cells were permeabilized with 0.1% Triton/PBS, for 10 min and blocked with 2% BSA for 30 min. The cells were then stained with rabbit α-GATA2 antibody (Santa Cruz Biotechnology, Inc) (1:1000) for 1 h followed by Alexa 594-conjugated goat anti-rabbit secondary antibody (Molecular Probes) (2 μg/ml) for 20 min. The slides were mounted in Vectashield® mounting medium with DAPI (Vector Laboratories, Inc). Cells without primary antibody served as negative controls. All incubations were performed at room temperature.

Western blot analysis

HL-60 cells carrying stably transduced 4HT-regulatable GATA2 (WT, p.Thr354Met, p.Thr355del and p.Leu359Val) were treated with and without 100 nM 4HT. After 24 h, the cells were harvested in RIPA buffer (50 mM Tris-Cl pH 7.6, 150 mM NaCl, 1% Triton X-100, 0.5% sodium deoxycholate, and 0.1% sodium dodecyl sulfate with protease inhibitor (cOmplete Mini EDTA free protease inhibitor tablets, Roche Diagnostics)). Samples were loaded onto the 10% acrylamide gels, electrophoresed and transferred onto Hybond-P PVDF membranes (Amersham). Membranes were probed with antibodies using standard techniques and visualised using ECL plus detection reagents (Amersham) on x-ray film (Amersham Hyperfilm™ MP).

Apoptosis Assays

HL-60 cells were stained for surface Annexin V and propidium iodide according to the manufacturer's protocol (#556547, Becton Dickinson).

Electromobility shift assay (EMSA) and EMSA-Western Blot

HEK293 cells were transfected with GATA2 WT or mutants using Lipofectamine™ 2000. After 24 h, nuclear extracts were prepared using a NE-PER® Nuclear and Cytoplasmic Extraction kit (Pierce) according to the manufacturer's protocol. Double stranded DNA oligonucleotides containing two GATA binding sites (Human TCRδ enhancer) or a single GATA binding site (GATA Consensus and Human GM-CSF-153 promoter) were synthesized (Supplementary Table 13). Each single stranded oligomer was labeled using a Biotin 3′ End DNA Labeling kit (Pierce) and annealed according to manufacturer's protocol. Electrophoretic mobility shift assays were performed using a modified protocol from Kumar *et al* 2008[35] and visualized using a Chemiluminescent Nucleic Acid Detection Module (Pierce) according to the manufacturer's protocol. Double stranded labeled probes (100 fmol) were incubated with 3 μg of nuclear extract for 20 min in 1x binding buffer containing 20 mM HEPES-KOH, pH 7.9, 100 mM KCl, 2 mM $MgCl_2$, 10 μM $ZnSO_4$, 10 mM 2

mercaptoethanol, 0.1% NP-40, 10% glycerol, 0.2 mM EDTA and 5 μg/ml sheared salmon sperm DNA. Polyclonal rabbit α-GATA2 (H-116) antibody (Cat. No. sc-9008; Santa Cruz Biotechnology, Inc) (1:100) was added to nuclear lysates for 20 min prior to addition of probe to demonstrate GATA2 as the binding protein. To assess the specificity of the binding, 200-fold excess of each unlabeled probe was used as competitor. The mixtures were resolved in 6% non-denaturing polyacrylamide gels made in 0.5x TGE buffer (12.5 mM Tris-HCl, pH 8.5, 85 mM glycine and 0.5 mM EDTA) and the electrophoresis was performed at 4°C. For EMSA-Western blots, the experiment was carried as described above, except that the shifted DNA oligonucleotides-protein complexes were transferred onto nitrocellulose membrane, instead of PVDF. The membrane was probed with monoclonal mouse α-GATA2 (CG2-96) antibody (Cat. No. sc-267; Santa Cruz Biotechnology, Inc) and detection was performed as mentioned above.

Determination of genes differentially expressed in the presence of GATA2 mutants

HL-60 cell lines were treated with 100 nM 4HT to turn on GATA2 WT and mutant protein expression. After 24 h, gene expression levels were determined by microarray(Supplementary Note).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank the families for their participation in our studies. Thanks also to Junia Melo for provision of CML cell lines and Raman Sharma for help with EMSA. Large scale Sanger sequencing was performed by the Australian Genome Research Facility, which was established through the Commonwealth-funded Major National Research Facilities program. This work was supported by grants from the National Health and Medical Research Council of Australia (program grants 257501 (HSS) and 219176 (HSS), fellowships 171601 (HSS) and 461204 (HSS), and a Dora Lush Postgraduate Award (CLC), Leukaemia Foundation of Australia (Grant in Aid to HSS, Postdoctoral fellowship to CLC), the Cancer Council of South Australia (HSS), and MedVet Pty Ltd (HSS and RDA) and NIH grant R01DK058161 (MSH).

References

1. Vardiman JW, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood. 2009; 114:937–951. [PubMed: 19357394]
2. Barzi A, Sekeres MA. Myelodysplastic syndromes: a practical approach to diagnosis and treatment. Cleve Clin J Med. 2010; 77:37–44. [PubMed: 20048028]
3. Owen C, Barnett M, Fitzgibbon J. Familial myelodysplasia and acute myeloid leukaemia--a review. Br J Haematol. 2008; 140:123–32. [PubMed: 18173751]
4. Zhang P, et al. Negative cross-talk between hematopoietic regulators: GATA proteins repress PU. 1. Proc Natl Acad Sci USA. 1999; 96:8705–8710. [PubMed: 10411939]
5. Crossley M, Merika M, Orkin SH. Self-association of the erythroid transcription factor GATA-1 mediated by its zinc finger domains. Mol Cell Biol. 1995; 15:2448–2456. [PubMed: 7739529]
6. Zhang SJ, et al. Gain-of-function mutation of GATA-2 in acute myeloid transformation of chronic myeloid leukemia. Proc Natl Acad Sci USA. 2008; 105:2076–2081. [PubMed: 18250304]
7. Yan XJ, et al. Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. Nat Genet. 2011; 43:309–15. [PubMed: 21399634]

8. Usary J, et al. Mutation of GATA3 in human breast tumors. Oncogene. 2004; 23:7669–7678. [PubMed: 15361840]
9. Owen CJ, et al. Five new pedigrees with inherited RUNX1 mutations causing familial platelet disorder with propensity to myeloid malignancy. Blood. 2008; 112:4639–4645. [PubMed: 18723428]
10. Carmichael CL, et al. Poor prognosis in familial acute myeloid leukaemia with combined biallelic CEBPA mutations and downstream events affecting the ATM, FLT3 and CDX2 genes. Br J Haematol. 2010; 150:382–385. [PubMed: 20456351]
11. Robinson CW, Norman JE, Cleland LG, Ford JH. Dermal necrosis and chromosome Iq abnormality in a man with a familial myeloproliferative disorder. Aust N Z J Med. 1983; 13:141–145. [PubMed: 6577833]
12. Horwitz M, Sabath DE, Smithson WA, Radich J. A family inheriting different subtypes of acute myelogenous leukemia. Am J Hematol. 1996; 52:295–304. [PubMed: 8701948]
13. Minelli A, et al. Familial partial monosomy 7 and myelodysplasia: different parental origin of the monosomy 7 suggests action of a mutator gene. Cancer Genet Cytogenet. 2001; 124:147–51. [PubMed: 11172908]
14. Michaud J, et al. Integrative analysis of RUNX1 downstream pathways and target genes. BMC Genomics. 2008; 9:363. [PubMed: 18671852]
15. Veitia RA. Dominant negative factors in health and disease. J Pathol. 2009; 218:409–18. [PubMed: 19544283]
16. Ernst T, et al. Transcription factor mutations in myelodysplastic/myeloproliferative neoplasms. Haematologica. 2010; 95:1473–1480. [PubMed: 20421268]
17. Matheny CJ, et al. Disease mutations in RUNX1 and RUNX2 create nonfunctional, dominant-negative, or hypomorphic alleles. EMBO J. 2007; 26:1163–1175. [PubMed: 17290219]
18. Kitajima K, et al. Redirecting differentiation of hematopoietic progenitors by a transcription factor, GATA-2. Blood. 2006; 107:1857–1863. [PubMed: 16254139]
19. Tsai FY, et al. An early haematopoietic defect in mice lacking the transcription factor GATA-2. Nature. 1994; 371:221–226. [PubMed: 8078582]
20. Zhang SJ, Shi JY. GATA-2 L359V mutation is solely associated with CML progression but not other hematological malignancies. Blood. 2008; 112:Abstract 1507.
21. Zhang SJ, Shi JY, Li JY. GATA-2 L359 V mutation is exclusively associated with CML progression but not other hematological malignancies and GATA-2 P250A is a novel single nucleotide polymorphism. Leuk Res. 2009; 33:1141–1143. [PubMed: 19304323]
22. Fadilah SA, Cheong SK, Roslan H, Rozie-Hanisa M, Yen GK. GATA-1 and GATA-2 gene expression is related to the severity of dysplasia in myelodysplastic syndrome. Leukemia. 2002; 16:1563–1565. [PubMed: 12145700]
23. Bullinger L, et al. Identification of acquired copy number alterations and uniparental disomies in cytogenetically normal acute myeloid leukemia using high-resolution single-nucleotide polymorphism analysis. Leukemia. 2010; 24:438–449. [PubMed: 20016533]
24. Wieser R, et al. Transcription factor GATA-2 gene is located near 3q21 breakpoints in myeloid leukemia. Biochem Biophys Res Commun. 2000; 273:239–245. [PubMed: 10873593]
25. Bullinger L, et al. Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. N Engl J Med. 2004; 350:1605–1616. [PubMed: 15084693]
26. Lahortiga I, et al. Molecular heterogeneity in AML/MDS patients with 3q21q26 rearrangements. Genes Chromosomes Cancer. 2004; 40:179–189. [PubMed: 15138998]
27. Slape C, et al. Retroviral insertional mutagenesis identifies genes that collaborate with NUP98-HOXD13 during leukemic transformation. Cancer Res. 2007; 67:5148–5155. [PubMed: 17545593]
28. Hsu AP, et al. Mutations in GATA2 are associated with the autosomal dominant and sporadic monocytopenia and mycobacterial infection (MonoMAC) syndrome. Blood. 2011
29. Mansour S, et al. Emberger syndrome-primary lymphedema with myelodysplasia: report of seven new cases. Am J Med Genet A. 2010; 152A:2287–96. [PubMed: 20803646]

30. Bigley V, et al. The human syndrome of dendritic cell, monocyte, B and NK lymphoid deficiency. J Exp Med. 2011; 208:227–34. [PubMed: 21242295]
31. Tsuzuki S, Towatari M, Saito H, Enver T. Potentiation of GATA-2 activity through interactions with the promyelocytic leukemia protein (PML) and the t(15;17)-generated PML-retinoic acid receptor alpha oncoprotein. Mol Cell Biol. 2000; 20:6276–86. [PubMed: 10938104]
32. Diessenbacher P, et al. NF-kappaB inhibition reveals differential mechanisms of TNF versus TRAIL-induced apoptosis upstream or at the level of caspase-8 activation independent of cIAP2. J Invest Dermatol. 2008; 128:1134–47. [PubMed: 17989734]
33. Callus BA, et al. Triggering of apoptosis by Puma is determined by the threshold set by prosurvival Bcl-2 family proteins. J Mol Biol. 2008; 384:313–23. [PubMed: 18835564]
34. Koldej R, Cmielewski P, Stocker A, Parsons DW, Anson DS. Optimisation of a multipartite human immunodeficiency virus based vector system; control of virus infectivity and large-scale production. J Gene Med. 2005; 7:1390–9. [PubMed: 16025547]
35. Kumar R, et al. CBFA2T3-ZNF652 corepressor complex regulates transcription of the E-box gene HEB. J Biol Chem. 2008; 283:19026–38. [PubMed: 18456661]

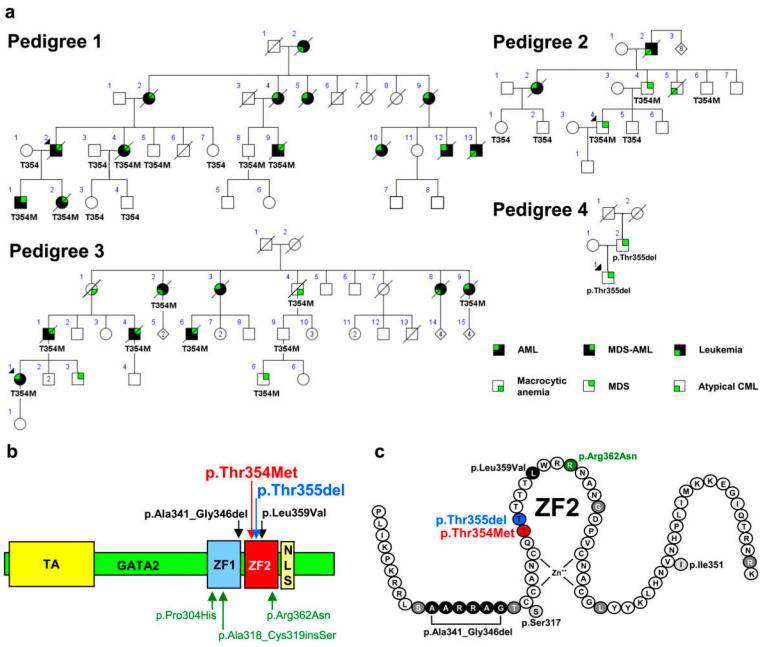

Figure 1. Identification of novel germline p.Thr354Met and p.Thr355del variants in the highly conserved zinc finger 2 domain of GATA2 that is associated with MDS-AML

a. Pedigrees containing the p.Thr354Met and p.Thr355del variants. One family from Australia (Pedigree 1) and two from the USA (Pedigrees 2 and 3) display the p.Thr354Met variant segregating with MDS-AML, and one USA family (Pedigree 4) contains a p.Thr355del variant that segregates with MDS. The genotype of tested individuals is shown; T354, (Thr354/Thr354); T354M, (Thr354/Met354). **b.** Domain structure of GATA2 showing positions of mutations. The positions of the p.Thr354Met, p.Thr355del, AML-M5[7] (green) and CML BC[6] (black) mutations are shown with respect to zinc finger (ZF) 1 and 2, transactivation domain (TA) and nuclear localization signal (NLS). **c.** Zinc finger 2 (ZF2) domain of GATA2 and GATA3 contains mutations associated with leukemia and breast cancer. The primary sequence is that of human GATA2 with the two alternative residues in GATA3 ZF2 shown (light grey with black letters). The position of p.Thr354Met and p.Thr355del is highlighted along with mutations found in GATA2 in AML-M5[7] (green) and CML BC[6] (black), and in GATA3 in breast cancer (summarized in [8]) (mutated residues in the corresponding GATA3 ZF2; grey with white letters).

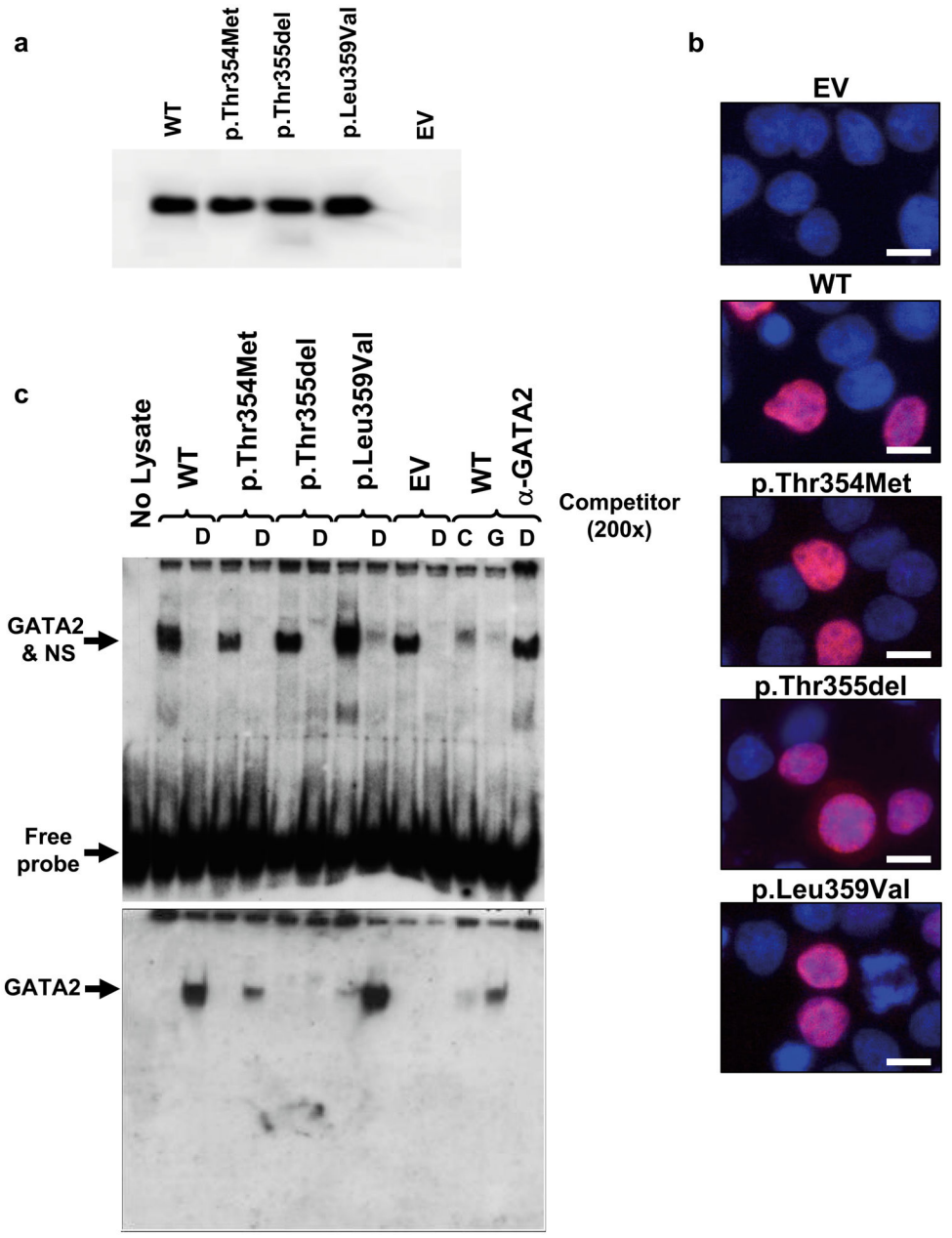

Figure 2. Subcellular localisation and DNA binding properties of GATA2 WT and mutants

HEK293 cells were transiently transfected with EV (pCMV-XL6 empty vector), WT, p.Thr354Met, p.Thr355del or p.Leu359Val and harvested after 24 h. **a.** Western blot analysis of GATA2 expression in nuclear lysates. Nuclear lysates were prepared and western blots performed, probing for GATA2. **b.** Cells were stained for GATA2 (red) and DAPI (blue). Scale bars, 10 μm. **c.** Electromobility shift assay (EMSA) of GATA2 WT and mutants. Nuclear lysates were prepared and bound to the TCRδ enhancer (contains GATA binding site) oligonucleotide in the absence or presence of 200-fold unlabeled competitor oligonucleotide (**D**, human TCRδ enhancer; **C**, GATA consensus; **G**, GM-CSF promoter). The probes were visualised using chemiluminescence (top panel). Note, GATA2 & NS

relates to a band that contains both GATA2 and a non-specific (NS) protein. To visualise GATA2 alone, an EMSA-western blot was performed probing with polyclonal α-GATA2 antibody (bottom panel), showing the level of binding of GATA2 WT and mutants. A neutralizing α-GATA2 antibody in the far right lane removes GATA2, but not the non-specific binding protein (NS) (top panel), and the specificity of GATA2 is confirmed in the bottom panel.

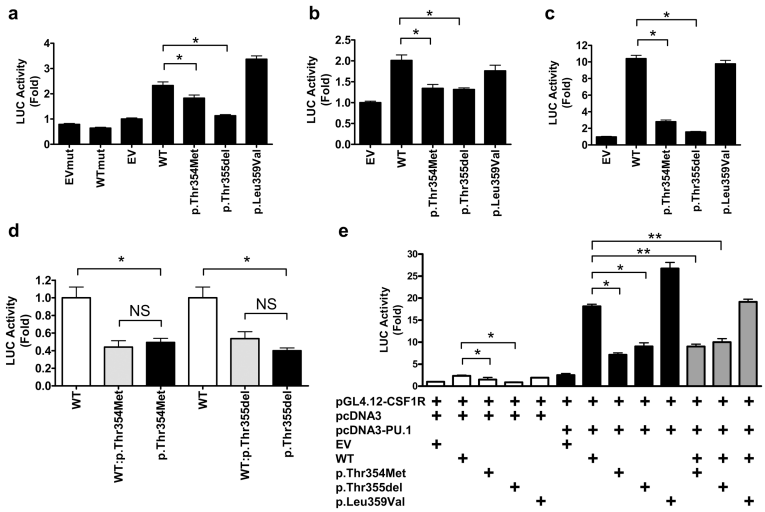

Figure 3. p.Thr354Met and p.Thr355del cause altered transactivation via target GATA2 response elements

p.Thr354Met and p.Thr355del act as a loss-of-function mutations on GATA2 target promoter and enhancer elements. HEK293 cells were cotransfected with 1) GATA2-responsive *CD34* (mut – *CD34* enhancer with GATA binding sites mutated[31]) (**a**) and *RUNX1* (**b**) enhancer elements linked to a LUC reporter, and 2) GATA2 (WT, p.Thr354Met, p.Thr355del or p.Leu359Val) expression constructs or pCMV6-XL6 empty vector (EV). Similarly, Cos-7 cells were cotransfected using *LYL1* promoter LUC as reporter (**c**). After 20 h, cells were harvested and luciferase assays performed and plotted as fold (mean ± s.e.m.) compared to EV control. Pairwise comparisons are shown (*p<0.05, n = 3). **d.** p.Thr354Met and p.Thr355del act as dominant negative mutations over WT GATA2. HEK293 cells were cotransfected with: 1) *CD34* enhancer-LUC reporter, and equivalent mole ratios of 2) WT to 3) p.Thr354Met or p.Thr355del. After 20 h, cells were harvested and luciferase assays performed. Pairwise comparisons are shown (*p<0.05; NS -not significant, n=3). **e.** p.Thr354Met has reduced ability to co-activate the *CSF1R* (*M-CSF-R*) promoter with PU.1. Cos-7 cells were cotransfected with 1) *CSF1R* promoter-LUC reporter, 2) PU.1 expression construct, and 3) WT, p.Thr354Met, p.Thr355del or p.Leu359Val expression constructs or EV. After 20 h, luciferase assays were performed and plotted as fold compared to EV. Pairwise comparisons are shown (*p<0.05, compared to WT plus PU.1; **p<0.05 compared to WT plus PU.1, but not significant when compared to p.Thr354Met or p.Thr355del plus PU.1, respectively). In all comparisons, a Student's t-test was used.

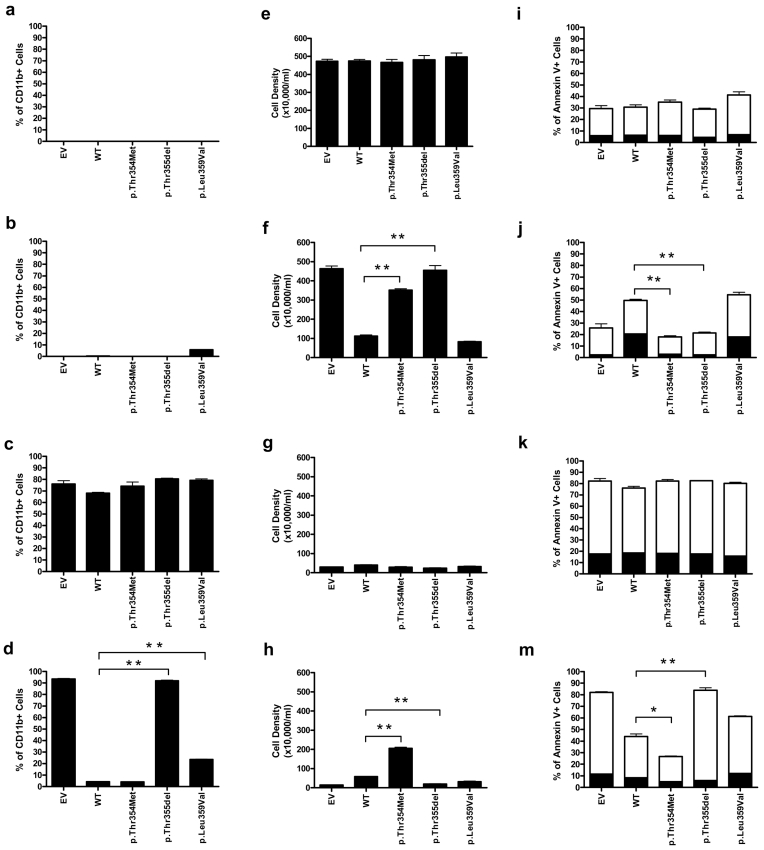

Figure 4. p.Thr354Met inhibits differentiation and apoptosis while allowing accumulation of cells in the presence of ATRA-induced differentiation

HL-60 cells carrying stably transduced 4HT-regulatable GATA2 (WT, p.Thr354Met, p.Thr355del and p.Leu359Val) or EV were treated with or without 30 nM 4HT for 24 h and then with or without 2 μM ATRA for 6 days. **a–d.** Differentiation of HL-60 cells into granulocytes. Differentiation was measured by FACS analysis for percentage of CD11b positive cells (mean ± s.e.m.) (see also Supplementary Fig. 7b). **e–h.** Cell numbers following differentiation. Cells were counted after 6 days (mean ± s.e.m.). **i–m.** Apoptosis following differentiation with ATRA. Cells were FACS analysed following staining with FITC anti-Annexin V and propidium iodide (PI). Annexin V^+, PI^- (black) or Annexin V^+, PI^+ (white). Indicative FACS plots (Supplementary Fig. 7c). **a,e,i.** −4HT, −ATRA; **b,f,j.** +4HT, −ATRA; **c,g,k.** −4HT, +ATRA; **d,h,m.** +4HT, +ATRA.(*p<0.05; **p<0.01, compared to WT). In all comparisons, a Student's t-test was used.