Germline mutations in *ETV6* are associated with thrombocytopenia, red cell macrocytosis and predisposition to lymphoblastic leukemia

Leila Noetzli[1,2,17], Richard W. Lo[3,4,17], Alisa B. Lee-Sherick[1], Michael Callaghan [5], Patrizia Noris[6], Anna Savoia[7] Madhvi Rajpurkar[5], Kenneth Jones[8], Katherine Gowan[8], Carlo Balduini[6], Alessandro Pecci[6], Chiara Gnan[7], Daniela De Rocco[7], Michael Doubek[9], Ling Li[3], Lily Lu[3], Richard Leung[3], Carolina Landolt-Marticorena[10], Stephen Hunger[1], Paula Heller[12], Arthur Gutierrez Hartman[13], Liang Xiayuan[14], Fred G. Pluthero[3], Jesse W.Rowley[15], Andrew S. Weyrich[15], Walter H. A. Kahr[3,4,16,17,18] , Christopher C. Porter[1,17,18] & Jorge Di Paola[1,2,17,18]

[1] Department of Pediatrics, University of Colorado Anschutz Medical Campus (AMC), Aurora, CO, USA
[2] Human Medical Genetics and Genomics Program, University of Colorado AMC, Aurora, CO, USA
[3] Program in Cell Biology, Research Institute, The Hospital for Sick Children, Toronto, ON, Canada
[4] Department of Biochemistry, University of Toronto, Toronto, ON, Canada
[5] Children's Hospital of Michigan, Wayne State University, Detroit, MI, USA
[6] Department of Internal Medicine, IRCCS Policlinico San Matteo Foundation, University of Pavia, Pavia, Italy
[7] Department of Medical Sciences, University of Trieste, Trieste, Italy and Institute for Maternal and Child Health- IRCCS Burlo Garofolo, Trieste, Italy
[8] Department of Biochemistry and Molecular Genetics, University of Colorado AMC, Aurora, CO, USA
[9] Department of Internal Medicine, Haematology/Oncology, University Hospital Brno, CZ
[10] Department of Medicine, University of Toronto, Division of Rheumatology University Health Network, Toronto Western Hospital, Toronto, ON, Canada
[12] Instituto de Investigaciones Medicas Alfredo Lanari, Universidad de Buenos Aires, Buenos Aires, Argentina
[13] Departments of Medicine, Biochemistry and Molecular Genetics. University of Colorado, AMC, Aurora, CO, USA
[14] Department of Pathology, University of Colorado, AMC, Aurora, CO, USA
[15] Department of Internal Medicine and the Molecular Medicine Program, University of Utah
[16] Department of Paediatrics, Division of Haematology/Oncology, University of Toronto and The Hospital for Sick Children, Toronto, ON, Canada
[17] These authors contributed equally to this work
[18] These authors are co-corresponding authors

Corresponding authors

Jorge Di Paola
Department of Pediatrics
Human Medical Genetics and Genomics Program
University of Colorado Anschutz Medical Campus
Research Complex 1, North Tower
12800 E. 19th Ave.
Mail Stop 8302 Room P18-4100
Aurora, CO 80045
Phone: 303 724 4000
Fax: 303 724 4015
Email: Jorge.dipaola@ucdenver.edu

Christopher C. Porter
Department of Pediatrics
University of Colorado Anschutz Medical Campus
Research Complex 1, North Tower
12800 E. 19th Ave.
Mail Stop 8302 Room P18-4101
Aurora, CO 80045
Phone: 303 724 4665
Fax: 303 724 4015
Email: chris.porter@ucdenver.edu

Walter HA Kahr
Departments of Paediatrics & Biochemistry,
Division of Haematology/Oncology,
Program in Cell Biology,
The Hospital for Sick Children,
University of Toronto,
555 University Avenue
Toronto, ON, M5G 1X8, CANADA
Tel: +1-416-813-7977
Fax: +1-416-813-5327
Email: walter.kahr@sickkids.ca

Some familial platelet disorders are associated with predisposition to leukemia, myelodysplastic syndrome (MDS) or dyserythropoietic anemia.[1,2] We have identified a family with autosomal dominant thrombocytopenia, high red blood cell mean corpuscular volume (MCV) and two occurrences of precursor B cell acute lymphoblastic leukemia (ALL). Whole exome sequencing identified a heterozygous single nucleotide change in *ETV6* (*Et*s *V*ariant Gene *6*), c.641C>T, encoding a p.Pro214Leu substitution in the central domain of the protein, which segregated exclusively with the thrombocytopenia and elevated MCV phenotype. Twenty-three additional families with a similar phenotype from a European consortium were screened, and two families were found to have mutations in *ETV6.* One family from the Czech Republic had the same p.Pro214Leu mutation with one individual developing B-cell ALL. The other family from Italy had a c.1252A>G nucleotide change that determined a p.Arg418Gly substitution in the DNA binding domain, as well as alternative splicing and exon skipping. Functional characterization of these mutations showed aberrant cytoplasmic localization of mutant and endogenous ETV6, decreased transcriptional repression by the ETV6 mutants, and altered maturation in hematopoietic cells cultured to become megakaryocytes. Our findings underscore a key role for ETV6 in platelet formation and leukemia predisposition.

ETV6 (also known as *TEL*) encodes a transcriptional repressor in the ETS family.[3] *ETV6* was initially identified as a tumor suppressor by involvement in somatic translocations in childhood leukemia, including *ETV6-RUNX1*, which often co-occur with a somatic mutation in the intact *ETV6* allele, suggesting that loss of function contributes to the development of leukemia.[4-6] Somatic *ETV6* mutations have also been described in patients with myelodysplastic syndrome (MDS) and T-cell leukemias,[7,8] but thus far germline mutations have not been described. *ETV6* also has a critical role in hematopoiesis, as demonstrated in animal models.[9,10] Erythroid/megakaryocytic conditional *Etv6* knockout mice are thrombocytopenic, and megakaryocyte colony formation is absent in homozygous knockout hematopoietic cells and decreased in heterozygotes, indicating *Etv6* involvement in thrombopoiesis.[11,12] Etv6 is also one of four transcription factors found sufficient to differentiate mouse fibroblasts into hematopoietic lineage cells.[13]

We present here three families with germline mutations in *ETV6* and defects in hematopoiesis. Affected members (n=5) in the original family from the United States (Family 1 Fig 1a) have variable thrombocytopenia (67,000-132,000 platelets/uL) and elevated red cell MCV (92.5-101.5 fL), suggesting a defect affecting megakaryocytic-erythroid precursors. Hematocrit and other hematologic indexes are within normal ranges (Supplementary Table 1). Platelets have normal mean volume and ultrastructure, with some elongated alpha granules (Supplementary Fig 1). Patients exhibit mild to moderate bleeding, and two developed precursor B cell ALL, at ages 3 (III-1) and 37 (II-7). Histopathological analysis of bone marrow from affected individuals without leukemia revealed small, hypolobulated megakaryocytes and abnormal red blood cell precursors (Fig 1b).

Whole exome sequencing of Family 1 revealed a heterozygous single nucleotide change in *ETV6,* which segregated completely with thrombocytopenia and high red blood cell MCV. This missense variant, c.641C>T, encodes a proline to leucine amino acid change (p.Pro214Leu) in the central domain of ETV6 (Fig 1c) that was not found in 1000 Genomes or dbSNP databases and was predicted to be possibly damaging by PolyPhen2. Twenty-three additional European families with autosomal dominant thrombocytopenia, high red blood cell MCV and increased incidence of leukemia were screened via Sanger sequencing of exons and exon-intron boundaries of *ETV6*. This approach found two families with germline *ETV6* mutations, one (Family 2, Fig 1a) with affected members having platelet counts of 44,000 -115,000 platelets/uL, MCV of 88-97 fL and ALL in individual I-2 at age 14, exhibited the identical c.641C>T mutation.

The other (Family 3, Fig 1a) with platelet counts of 99,000-101,000 platelets/uL and MCV of 93-98 fL, but no malignancies had a mutation in the DNA- binding domain (c.1252A>G, p.Arg418Gly) not observed in 1000 Genomes and predicted to be highly damaging by PolyPhen2. Sequence alignment demonstrated that both mutations affect amino acids that are highly conserved across multiple species (Supplementary Information).

The c.1252A>G mutation is located in the last codon of exon 7, which is split with exon 8, raising the possibility of this variant disrupting a splice site. To test this possibility RNA was isolated from peripheral blood cells from two individuals with the c.1252A>G mutation. RT-PCR detected two different transcripts, one of expected size (386 bp) and another of 285 bp, indicating an alternatively-spliced product (Supplementary Fig 2). Sequencing of the 285bp product revealed skipping of exon 7. This c.1153_1253del mutation is predicted to cause a partial deletion of the putative DNA binding domain (aa 385-418, p.385_418del) and a subsequent p.Asn385Valfs*7 frameshift alteration resulting in a premature stop codon raising the possibility of a truncated protein. Although this truncated protein was expressed in transfection assays in HEK293T cells (Supplementary Fig 3) it was not detected in patients' platelets (Supplementary Fig 4), suggesting it is not functional in megakaryocytes. The alternative splicing did not affect all mutant RNA, since sequencing of the 386 bp RT-PCR product, as well as plasmids where this product was cloned, showed the presence of the G nucleotide in addition to the wild type A, indicating that the p.Arg418Gly form of ETV6 is likely to be expressed in patients from this family. This type of genetic aberration, in which an exonic mutation creates an alternative splice site as well as an amino acid change, has been observed in other hematological disorders such as Hemoglobin E disease.[14]

ETV6 is a 57kD protein with 452 amino acids and three functional domains: N-terminal pointed (PNT), central regulatory and C-terminal DNA-binding (ETS; Fig 1c). Nuclear localization and transcriptional repression activity of ETV6 require homodimerization via the pointed domain.[15] ETV6 modulates the activity of other ETS transcription factors such as FLI1, where hemizygous deletions cause Paris-Trousseau related thrombocytopenia,[16,17] and several other known ETV6 interaction partners that are present in platelets,[18,19] and presumably megakaryocytes. The function of the central domain is not well understood, but it has been shown to undergo posttranslational modifications and to be essential for the repressive function of ETV6 in *in vitro* reporter assays.[20]

To investigate the effect of the p.Pro214Leu, p.Arg418Gly, and p.385_418del mutations on ETV6 transcriptional repression we used an *in vitro* reporter assay to measure luciferase expression induced by the promoter of the known ETV6 target *stromelysin-1*.[8,21] As expected, wild type ETV6 (WT) repressed luciferase expression compared to empty vector. Expression of each of the three mutant forms of ETV6 resulted in less transcriptional repression when expressed alone or WT ETV6 (Figure 1d). Protein expression of WT and all ETV6 mutants in transfected cells were determined to be equivalent by immunoblot, suggesting that these mutations do not influence protein stability (Supplementary Fig 3).

Thrombopoiesis involves a complex sequence of cellular events in bone marrow megakaryocytes culminating with the elaboration of proplatelet extensions that release platelets into the circulation. Approximately 10^{11} platelets must be produced daily to maintain normal concentrations of 150-400 x 10^9 platelets per liter of human blood.[22] To determine the effect of P214L and R418G ETV6 on megakaryocyte differentiation human CD34+ cells were transduced with lentivirus containing WT, P214L or R418G ETV6 alleles and cultured with thrombopoietin to support megakaryocyte development. P214L and R418G ETV6-transduced cells showed delayed/decreased maturation when compared to those transduced with WT ETV6, indicated by increased numbers of small, immature megakaryocytes (similar to the abnormalities observed in the bone marrow of affected individuals, Fig 1b) and decreased generation of proplatelets (Fig 2 a-d). These findings indicate that these mutations affect megakaryocyte development and likely platelet production.

Since ETV6 requires dimerization to exert transcriptional repression, it is possible that the P214L, R418G and p.385-418del mutants may affect this interaction in a dominant negative manner. All mutants were shown to dimerize with wild type ETV6, demonstrated by concomitant “pull down” of both differentially tagged his-ETV6 WT and myc-ETV6 mutants (Supplementary Fig 5). The intracellular distribution of protein produced by transduced P214L or R418G and WT forms of ETV6 in megakaryocytes was examined by immunofluorescence staining for their *myc* tag, which showed expected nuclear localization for WT ETV6, while P214L and R418G ETV6 were concentrated in the cytoplasm (Supplementary Figure 6). Staining for ETV6 (Figure 3) showed this protein concentrated in the nuclei of an untransduced control and WT-transduced cells, and largely in the cytoplasm of cells expressing mutant ETV6 (see also Videos 1-4). These results indicate that the mutant forms of ETV6 mislocalize inside cells, and by dimerizing with

endogenous WT protein, may also prevent nuclear localization, further compromising ETV6 repression activity.

Several cases of leukemia involving somatic *ETV6* alterations have been associated with loss of *ETV6* function.[4-6] In order to identify additional genetic aberrations contributing to leukemogenesis, whole exome and RNA sequencing were performed on leukemia blast cells from time of diagnosis, and from bone marrow cells after remission in the same patient (III-1, Family 1). RNA sequencing revealed that the wild type *ETV6* allele was structurally normal and that wild type and p.P214L transcripts were expressed at equal amounts in the bone marrow samples. Several genetic variants and a novel gene fusion between *PAX5* and *SHB* were found in the leukemia sample and were not present in the remission sample (Supplementary Table 2). These variants may represent candidate cooperative mutations with the germline *ETV6* mutation in the development of leukemia.[23, 24]

In summary, we present novel germline mutations in *ETV6,* a gene known to be involved in human leukemogenesis and murine and zebra fish hematopoiesis, in three families with thrombocytopenia and high red cell MCV and leukemia predisposition in two of the three families. The p.Pro214Leu and p.Arg418Gly mutations partially disrupt ETV6 transcriptional repression *in vitro* and cause aberrant cytoplasmic localization in megakaryocytes, suggesting a dominant negative phenotype due to the requirement of ETV6 dimerization and nuclear localization for transcriptional repression. The p.Pro214Leu and the p.Arg418Gly mutations also impair development of megakaryocytes in culture. The mutations in these families reveal the implications of viable germline mutations in *ETV6,* and provide insights into the importance of ETV6 for platelet biology and tumor suppressor activity.

Acknowledgments
We are grateful to the families studied for their contribution to this project. We are also grateful to Tamim Shaikh, Rich Spritz and Jeff Murray for their insightful comments. This work was supported by the Postle Family Chair in Pediatric Cancer and Blood Disorders (JDP), NIH HL112311 (ASW), GM103806 (JWR). WHAK was supported by operating grants from the Canadian Institutes of Health Research (CIHR; MOP-81208 and MOP-259952). PN and AS were supported by grant GGP13082 from Telethon Foundation.

References

1. Song, W.J. *et al.* Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. *Nat Genet* **23**, 166-75 (1999).
2. Nichols, K.E. *et al.* Familial dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA1. *Nat Genet* **24**, 266-70 (2000).
3. Kar, A. & Gutierrez-Hartmann, A. Molecular mechanisms of ETS transcription factor-mediated tumorigenesis. *Crit Rev Biochem Mol Biol* **48**, 522-43 (2013).
4. Romana, S.P. *et al.* Deletion of the short arm of chromosome 12 is a secondary event in acute lymphoblastic leukemia with t(12;21). *Leukemia* **10**, 167-70 (1996).
5. Patel, N. *et al.* Expression profile of wild-type ETV6 in childhood acute leukaemia. *Br J Haematol* **122**, 94-8 (2003).
6. Barjesteh van Waalwijk van Doorn-Khosrovani, S. *et al.* Somatic heterozygous mutations in ETV6 (TEL) and frequent absence of ETV6 protein in acute myeloid leukemia. *Oncogene* **24**, 4129-37 (2005).
7. Bejar, R. *et al.* Clinical effect of point mutations in myelodysplastic syndromes. *N Engl J Med* **364**, 2496-506 (2011).
8. Van Vlierberghe, P. *et al.* ETV6 mutations in early immature human T cell leukemias. *J Exp Med* **208**, 2571-9 (2011).
9. Rasighaemi, P., Onnebo, S.M., Liongue, C. & Ward, A.C. ETV6 (TEL1) regulates embryonic hematopoiesis in zebrafish. *Haematologica* (2014).
10. Wang, L.C. *et al.* Yolk sac angiogenic defect and intra-embryonic apoptosis in mice lacking the Ets-related factor TEL. *EMBO J* **16**, 4374-83 (1997).
11. Wang, L.C. *et al.* The TEL/ETV6 gene is required specifically for hematopoiesis in the bone marrow. *Genes Dev* **12**, 2392-402 (1998).
12. Hock, H. *et al.* Tel/Etv6 is an essential and selective regulator of adult hematopoietic stem cell survival. *Genes Dev* **18**, 2336-41 (2004).
13. Pereira, C.F. *et al.* Induction of a hemogenic program in mouse fibroblasts. *Cell Stem Cell* **13**, 205-18 (2013).
14. Orkin, S.H. *et al.* Abnormal RNA processing due to the exon mutation of beta E-globin gene. *Nature* **300**, 768-9 (1982).
15. Green, S.M., Coyne, H.J., 3rd, McIntosh, L.P. & Graves, B.J. DNA binding by the ETS protein TEL (ETV6) is regulated by autoinhibition and self-association. *J Biol Chem* **285**, 18496-504 (2010).
16. Kwiatkowski, B.A. *et al.* The ets family member Tel binds to the Fli-1 oncoprotein and inhibits its transcriptional activity. *J Biol Chem* **273**, 17525-30 (1998).
17. Raslova, H. *et al.* FLI1 monoallelic expression combined with its hemizygous loss underlies Paris-Trousseau/Jacobsen thrombopenia. *J Clin Invest* **114**, 77-84 (2004).
18. Million, R.P., Harakawa, N., Roumiantsev, S., Varticovski, L. & Van Etten, R.A. A direct binding site for Grb2 contributes to transformation and leukemogenesis by the Tel-Abl (ETV6-Abl) tyrosine kinase. *Mol Cell Biol* **24**, 4685-95 (2004).
19. Roukens, M.G., Alloul-Ramdhani, M., Moghadasi, S., Op den Brouw, M. & Baker, D.A. Downregulation of vertebrate Tel (ETV6) and Drosophila Yan is facilitated by an evolutionarily conserved mechanism of F-box-mediated ubiquitination. *Mol Cell Biol* **28**, 4394-406 (2008).
20. Lopez, R.G. *et al.* TEL is a sequence-specific transcriptional repressor. *J Biol Chem* **274**, 30132-8 (1999).
21. Fenrick, R. *et al.* TEL, a putative tumor suppressor, modulates cell growth and cell morphology of ras-transformed cells while repressing the transcription of stromelysin-1. *Mol Cell Biol* **20**, 5828-39 (2000).

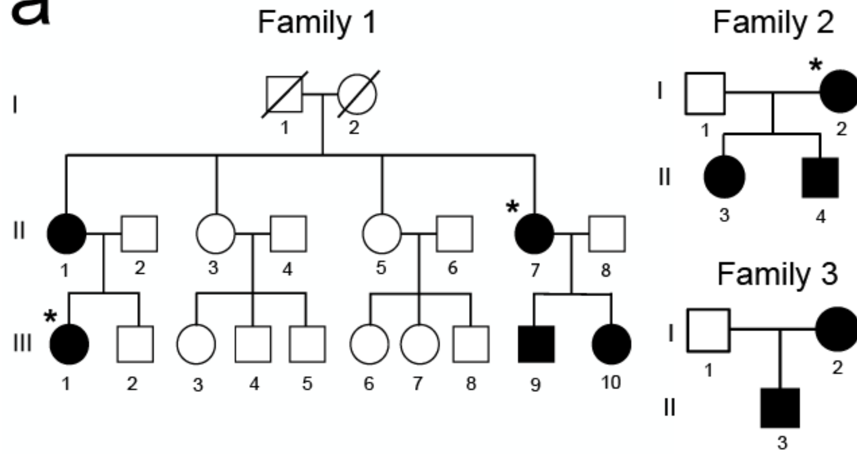
22. Machlus, K.R. & Italiano, J.E., Jr. The incredible journey: From megakaryocyte development to platelet formation. *J Cell Biol* **201**, 785-96 (2013).
23. Mullighan, C.G. *et al.* Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* **446**, 758-64 (2007).
24. Strehl, S., Konig, M., Dworzak, M.N., Kalwak, K. & Haas, O.A. PAX5/ETV6 fusion defines cytogenetic entity dic(9;12)(p13;p13). *Leukemia* **17**, 1121-3 (2003).

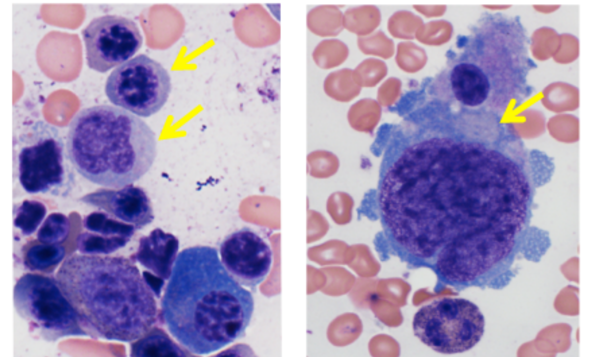
Figure Legends

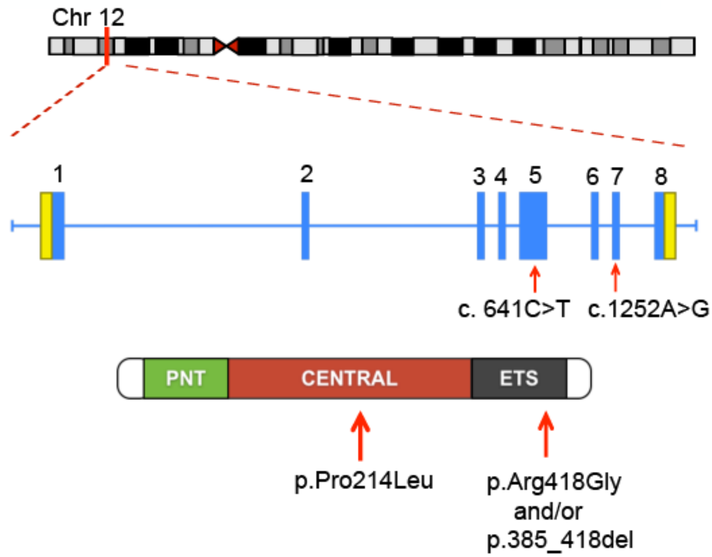
Figure 1. Mutation analysis of *ETV6.* (a) Pedigrees for three affected families. Solid symbols represent affected individuals with thrombocytopenia and high red cell mean corpuscular volume. Asterisks represent individuals that developed B cell leukemia. (Supplementary Table 1 shows all CBC values for affected and non-affected individuals). **(b)** Representative images of a bone marrow aspirate for one of the affected individuals without leukemia that shows mild dyserythropoiesis in the left panel (yellow arrows) and an immature hypolobulated megakaryocyte (yellow arrow) on the right panel. **(c)** Schematic of *ETV6,* which is composed of 8 exons encoding untranslated regions (yellow) and protein coding sequences (blue). Two mutations, c.641C>T in exon 5 and c.1252A>G in exon 7 are depicted. In the lower section of the panel a schematic of ETV6 is represented with its different domains, pointed (PNT), central and ETS (which is the DNA-binding domain) and the location of the amino acid changes. **(d)** Effects of *ETV6* mutant alleles (p.Pro214Leu, p.Arg418Gly, and p.385_418del) in the activity of an ETV6-responsive reporter plasmid (pGL2-754TR) when expressed alone or with the wild type allele. Luciferase activity, represented here by fold repression, is shown relative to empty vector and normalized using an internal control plasmid expressing *Renilla* luciferase. Wild type ETV6 (WT) repressed luciferase expression by approximately 7.5 fold compared to empty vector, whereas p.P214L ETV6, p.R418G, and p.385_418del repressed transcription 4.5, 1.5, and 1.5 fold, respectively ($p<0.0001$, one way ANOVA) (Figure 1d). Co-expression of WT and mutant proteins caused intermediate repression of the *stromelysin-1* promoter. Experiments were done in triplicates and repeated at least 4 times.

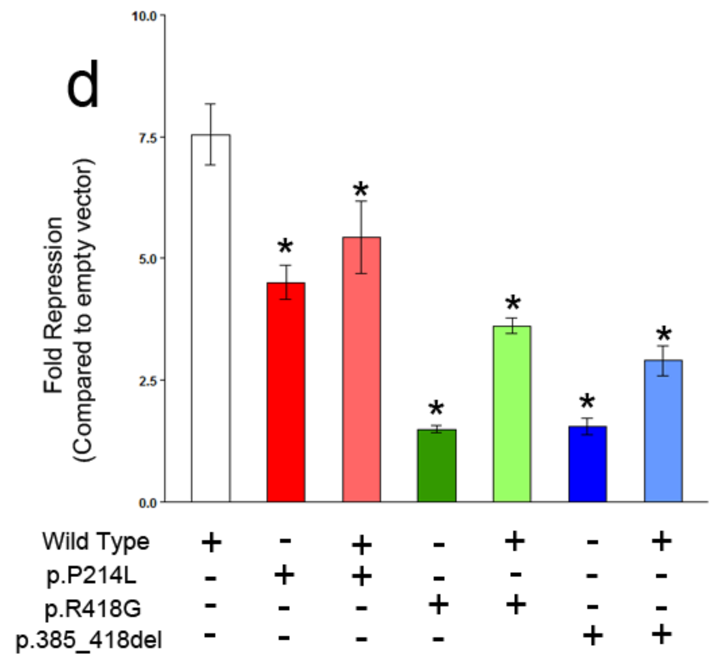
Figure 2. Abnormal development of day 12 cultured megakaryocytes expressing mutant ETV6. Control cells and those transduced with lentiviral myc-tagged forms of ETV6 ($ETV6^{WT}$, $ETV6^{P214L}$ and $ETV6^{R418G}$) were assessed via immunofluorescence microscopy imaging. (a-c) Megakaryocytes (control cells shown) were identified via expression of CD61 (violet) and VWF (red) and staged by diameter: <15 μm (a), >15μm (b), or the presence of proplatelets (c). (d) Population distributions showed no significant differences between control and $ETV6^{WT}$-transduced cells, while $ETV6^{P214L}$-transduced megakaryocytes showed a significantly higher proportion (*$P<0.05$, 2-tailed t-test) of cells in the earlier developmental stage (<15 μm) and fewer in the mature proplatelet-forming stage (control: n=7, WT: n=3, P214L: n=3, R418G: n=4 cultures with >300 cells for each). Images for control cells and those transduced with WT, P214L and R418G forms of ETV6 are shown in Supplementary Fig 6.

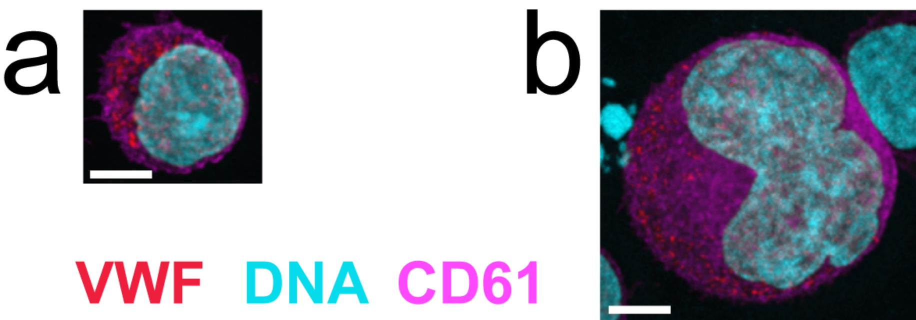
Figure 3. **Aberrant cytoplasmic localization of ETV6 in cultured megakaryocytes transduced with ETV6 mutants.** Comparison via immunofluorescence microscopy of ETV6 localization in Day 12 cultured control megakaryocytes and those transduced with lentiviral wild-type ($ETV6^{WT}$) or mutant ($ETV6^{P214L}$, $ETV6^{R418G}$) myc-tagged forms of ETV6. Confocal Z-sections of representative mature cells (>15μm) are shown stained for DNA (blue), ETV6 (both endogenous and myc-tagged; green) and myc (red). In control cells endogenous ETV6 is concentrated in the nucleus, as is endogenous and myc-tagged WT ETV6 in cells transduced with $ETV6^{WT}$. In contrast, cells transduced with mutant myc-tagged ETV6 ($ETV6^{P214L}$, $ETV6^{R418G}$) show this protein concentrated in the cytoplasm with little visible in the nucleus. Scale bars = 5 μm. See also Videos 1-4.

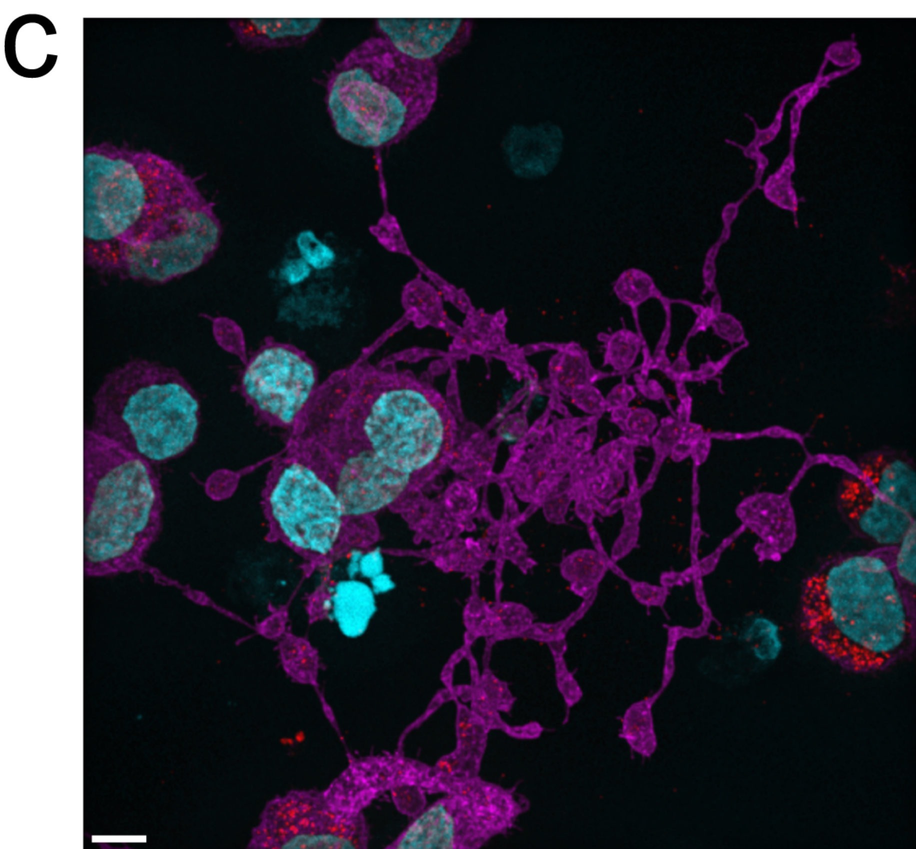

a
Family 1
Family 2
Family 3
I
II
III

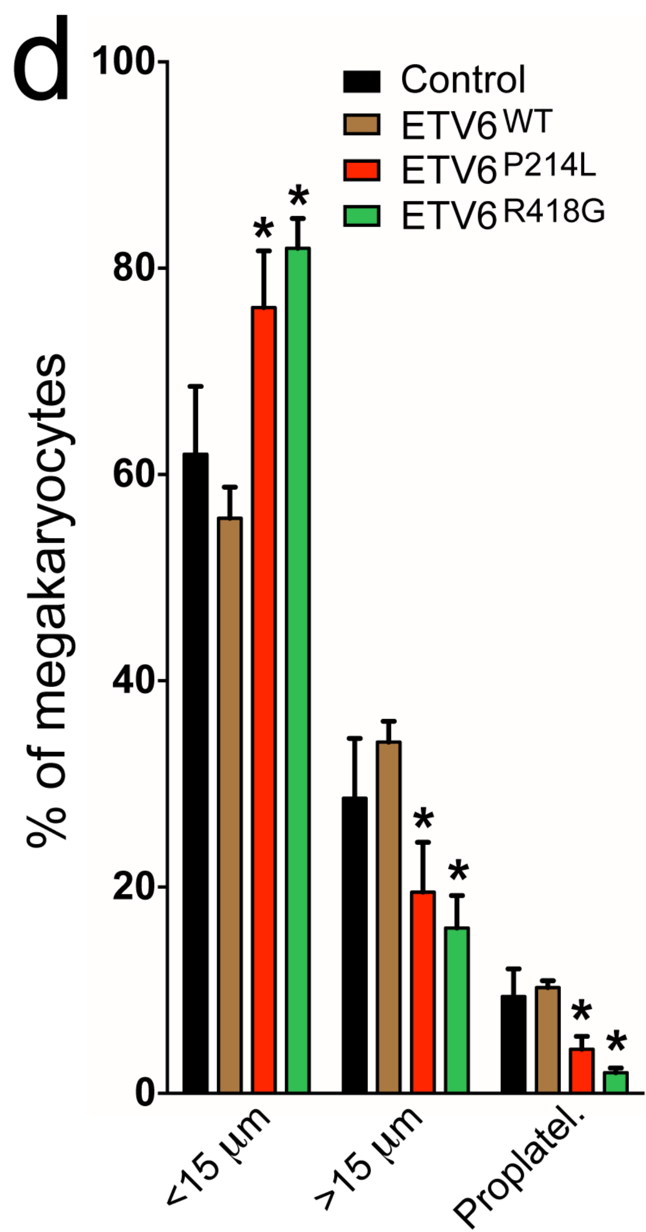

b

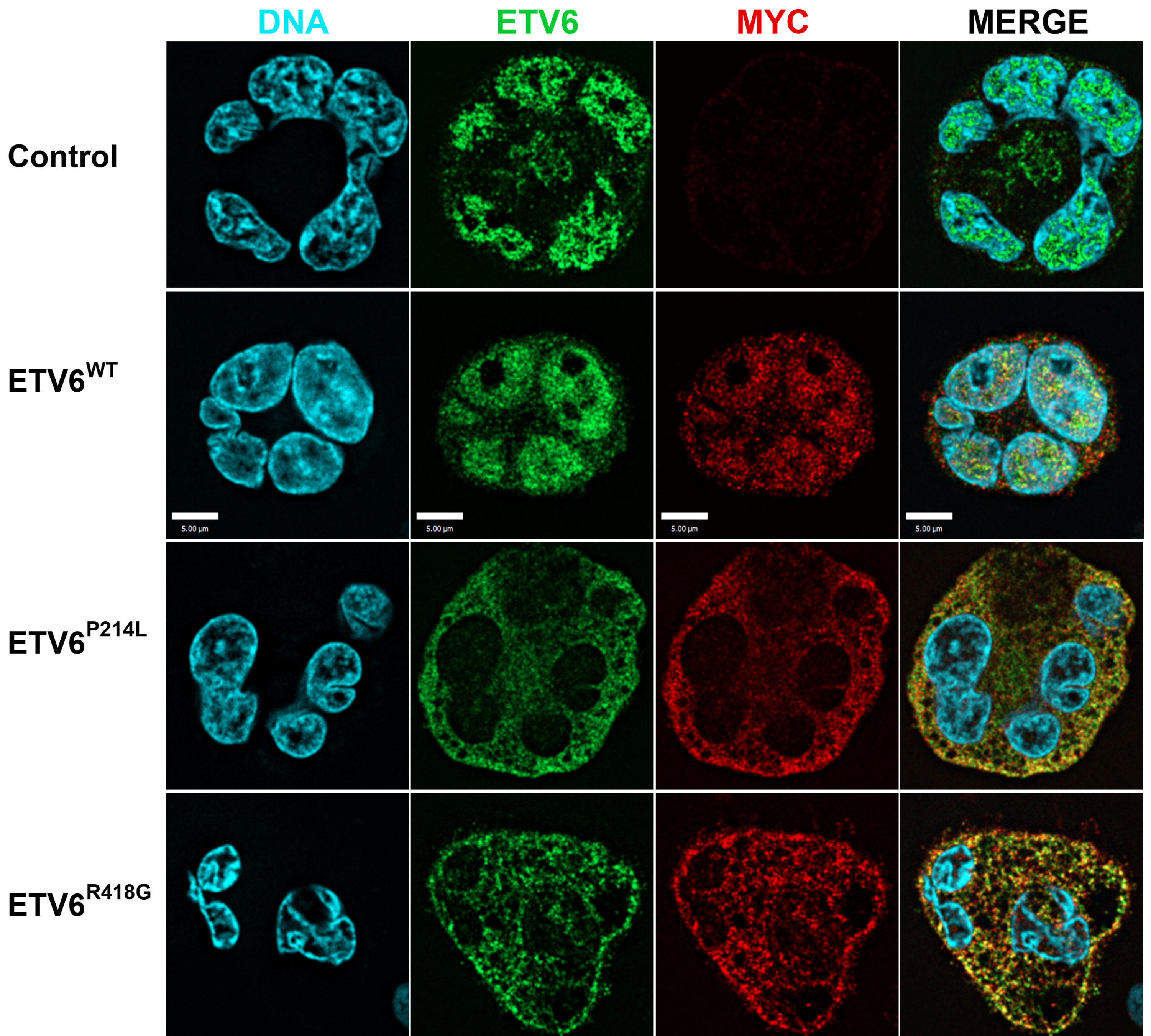

c
Chr 12
c. 641C>T
c.1252A>G
PNT
CENTRAL
ETS
p.Pro214Leu
p.Arg418Gly
and/or
p.385_418del


d
Fold Repression
(Compared to empty vector)
Wild Type
p.P214L
p.R418G
p.385_418del


a
b
VWF DNA CD61


c


d
Control
ETV6 WT
ETV6 P214L
ETV6 R418G
% of megakaryocytes
0
20
40
60
80
100
<15 µm
>15 µm
Proplatel.



DNA
ETV6
MYC
MERGE
Control
ETV6WT
ETV6P214L
ETV6R418G
5.00 µm
5.00 µm
5.00 µm
5.00 µm