OPEN ACCESS

University of
BRISTOL

Turro, E., Greene, D., Wijgaerts, A., Thys, C., Lentaigne, C., Bariana, T. K.,
... Ouwehand, W. H. (2016). A dominant gain-of-function mutation in
universal tyrosine kinase SRC causes thrombocytopenia, myelofibrosis,
bleeding, and bone pathologies. Science Translational Medicine, 8(328),
[328ra30]. DOI: 10.1126/scitranslmed.aad7666

Peer reviewed version

Link to published version (if available):
10.1126/scitranslmed.aad7666

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research
General rights

A dominant gain-of-function mutation in universal tyrosine kinase *SRC* causes thrombocytopenia, myelofibrosis, bleeding and bone pathologies

Ernest Turro[1-4], Daniel Greene[1,3,4], Anouck Wijgaerts[5], Chantal Thys[5], Claire Lentaigne[6,7], Tadbir K Bariana[8,9], Sarah K Westbury[10], Anne M Kelly[1,2], Dominik Selleslag[11], Jonathan C Stephens[1,2,4], Sofia Papadia[1,4], Ilenia Simeoni[1,4], Christopher J Penkett[1,4], Sofie Ashford[1,4], Antony Attwood[1,2,4], Steve Austin[12], Tamam Bakchoul[13], Peter Collins[14], Sri V V Deevi[1,4], Rémi Favier[15], Myrto Kostadima[1,2], Michele P Lambert[16,17], Mary Mathias[18], Carolyn M Millar[6,7], Kathelijne Peerlinck[5], David J Perry[19], Sol Schulman[20], Deborah Whitehorn[1,2], Christine Wittevrongel[5], BRIDGE-BPD Consortium[#], Marc De Maeyer[21], Augusto Rendon[1,22], Keith Gomez[8,9], Wendy N Erber[23], Andrew D Mumford[10,24], Paquita Nurden[25], Kathleen Stirrups[1,4], John R Bradley[4,26], F Lucy Raymond[4,27], Michael A Laffan[6,7], Chris Van Geet[5], Sylvia Richardson[3], Kathleen Freson[5,*,‡] & Willem H Ouwehand[1,2,4,28,*]

Affiliations

[1] Department of Haematology, University of Cambridge, Cambridge Biomedical Campus, Cambridge, United Kingdom. [2] NHS Blood and Transplant, Cambridge Biomedical Campus, Cambridge, United Kingdom. [3] Medical Research Council Biostatistics Unit, Cambridge Institute of Public Health, Cambridge Biomedical Campus, Cambridge, United Kingdom. [4] NIHR BioResource - Rare Diseases, Cambridge University Hospitals, Cambridge Biomedical Campus, Cambridge, United Kingdom. [5] Department of Cardiovascular Sciences, Center for Molecular and Vascular Biology, University of Leuven, Belgium. [6] Centre for Haematology, Hammersmith Campus, Imperial College Academic Health Sciences Centre, Imperial College London, London, United Kingdom. [7] Imperial College Healthcare NHS Trust, Du Cane Road, London, United Kingdom. [8] Department of Haematology, University College London Cancer Institute, London, United Kingdom. [9] The Katharine Dormandy Haemophilia Centre and Thrombosis Unit, Royal Free London NHS Foundation Trust, London, United Kingdom. [10] School of Clinical Sciences, University of Bristol, United Kingdom. [11] Academisch Ziekenhuis Sint Jan, Brugge-Oostende, Belgium. [12] Department of Haematology, Guy's and St Thomas' NHS Foundation Trust, London, United Kingdom. [13] Institute for Immunology and Transfusion Medicine, Universitätsmedizin Greifswald, Germany. [14] Arthur Bloom Haemophilia Centre, Institute of Infection and Immunity, School of Medicine, Cardiff University, United Kingdom. [15] Assistance Publique -Hôpitaux de Paris, Armand Trousseau Children Hospital, Paris; Inserm U1170, Villejuif, France. [16] Division of Hematology, Children's Hospital of Philadelphia, Philadelphia, United States of America. [17] Department of Pediatrics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, United States of America. [18] Department of Haematology, Great Ormond Street Hospital for Children NHS Foundation Trust, London, United Kingdom. [19] Department of Haematology, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge Biomedical Campus, Cambridge, United Kingdom. [20] Beth Israel Deaconess Medical Centre, Harvard Medical School, Boston, United

States of America. [21] Biochemistry, Molecular and Structural Biology Section, University of Leuven, Belgium. [22] Genomics England Ltd, London, United Kingdom. [23] Pathology and Laboratory Medicine, University of Western Australia, Crawley, Western Australia, Australia. [24] School of Cellular and Molecular Medicine, University of Bristol, Bristol, United Kingdom. [25] Institut Hospitalo-Universitaire LIRYC, PTIB, Hôpital Xavier Arnozan, Pessac, France. [26] Research & Development, Cambridge University Hospitals NHS Foundation Trust, Cambridge, United Kingdom. [27] Department of Medical Genetics, Cambridge Institute for Medical Research, University of Cambridge, Cambridge, United Kingdom. [28] Human Genetics, Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, United Kingdom.

* Authors contributed equally to the study
‡ Corresponding author (kathleen.freson@med.kuleuven.be)
# A list of additional members of the BRIDGE-BPD Consortium is provided in the supplementary materials.

SUMMARY: Increased podosome formation due to E527K hyperactive SRC results in thrombocytopenia, myelofibrosis, bleeding and bone pathologies

ABSTRACT (237/250)
The Src family kinase (SFK) member SRC is a major target in drug development because it is activated in many human cancers, yet deleterious *SRC* germline mutations have not been reported. We used genome sequencing and Human Phenotype Ontology patient coding to identify a gain-of-function mutation in *SRC* causing thrombocytopenia, myelofibrosis, bleeding and bone pathologies in nine cases. Modeling of the E527K substitution predicts loss of SRC's self-inhibitory capacity, which we confirmed with *in vitro* studies showing increased SRC kinase activity and enhanced Tyr419 phosphorylation in COS7 cells overexpressing E527K SRC. The active form of SRC predominates in patients' platelets, resulting in enhanced overall tyrosine phosphorylation. Patients with myelofibrosis have hypercellular bone marrow with trilineage dysplasia and their stem cells grown *in vitro* form more myeloid and megakaryocyte (MK) colonies than control cells. These MKs generate platelets which are dysmorphic, low in number, highly variable in size and with a paucity of α-granules. Over-active SRC in patient-derived MKs causes a reduction in proplatelet formation, which can be rescued by SRC kinase inhibition. Stem cells transduced with lentiviral E527K SRC form MKs with a similar defect and enhanced tyrosine phosphorylation levels. Patient-derived and E527K-transduced MKs show Y419 SRC-positive stained podosomes that induce altered actin organization. Expression of mutated *src* in zebrafish recapitulates patients' blood and bone phenotypes. Similar studies of platelets and MKs may reveal the mechanism underlying the severe bleeding frequently observed in cancer patients treated with next-generation SFK inhibitors.

INTRODUCTION

The Rous sarcoma virus (RSV), via its v-*src* viral gene lacking the codon for the tyrosine(Y)-527 self-inactivation residue, was described in the mid-20th century as causing tumor growth in chickens (*1*, *2*). v-Src increased phosphotyrosine levels in cellular proteins within RSV-transformed cells (*3*) and induced actin-based dynamic protrusions of the plasma membrane, named podosomes and invadosomes (*4,5*). The highly similar proto-oncogene tyrosine-protein kinase SRC (or c-Src) is a non-receptor tyrosine kinase protein that is encoded by the *SRC* gene. A somatic truncating mutation at *SRC* residue 531 has been detected in human colon cancer and also results in SRC activation (*6*). Of all Src family kinases (SFKs), SRC is the most frequently implicated in cancer and various SFK inhibitors are in clinical trials (*7*, *8*). Dasatinib, Bosutinib and Ponatinib were developed as next-generation inhibitors against Bcr-Abl to treat chronic myeloid leukemia (CML) and other targets (including SFKs) (*9*). Although these drugs are not SRC-specific, it is remarkable that approximately 10-40% of CML cases treated develop hematological side-effects, including severe bleeding which is disproportionate to the reduction in platelet counts (*9, 10*). Pathogenic germline mutations in *SRC*, however, have never been described in humans.

The highest SRC protein levels are found in platelets (*11*) and brain (*12*), with only moderate expression in other tissues. Remarkably, *Src* knockout mice have no detectable brain phenotype, no bleeding symptoms and a normal platelet count (*13*). Most homozygous mice die within the first weeks of birth and those which survive present only with impaired osteoclast function, osteopetrosis and a failure of incisors to erupt. Over the last three decades, several activation pathways in platelets have been studied that are regulated by different SFKs. In both mice and humans, platelet signalling events downstream of platelet integrins, G-protein coupled receptors for agonists and the receptor for VWF are described as being dependent on SRC (*14*). Here, we show that a germline gain-of-function *SRC* mutation leads to thrombocytopenia, myelofibrosis, bleeding, platelet dysfunction with abnormal α-granules and bone pathologies.

RESULTS

Statistical analysis of genetic and phenotypic data

We studied a three-generation pedigree with a dominant inheritance of bleeding with fatal consequences for one of the nine cases. Microscopic inspection of the blood film showed a reduced number of platelets of which about 10-30% have a grayish appearance due to a lack of α-granules (Fig. 1A, table 1). Five cases presented prematurely with myelofibrosis accompanied by enlargement of their spleens (table 1). Three cases underwent splenectomy, which did not correct the platelet count or their atypical morphology. Cases in this pedigree also have extensive bone pathologies with edentulism before the third decade and mild dysmorphism characterised by large forehead, ocular hypotelorism, deep-set eyes and a wide-nostriled nose (table 1). The clinical and laboratory phenotypes of three cases were coded with Human Phenotype Ontology (HPO) terms (*15*), including terms relating to platelet morphology (Fig. 1C). Mutations in known thrombocytopenia genes were absent, prompting us to subject DNA samples from two cases (Fig. 1A) to genome sequencing. After filtering by alternate allele

frequency in reference collections and selecting variants predicted to affect translation, we determined that these two cases shared 67 plausible causal variants in 67 candidate genes. Building on recent approaches for HPO-based differential diagnosis (*16*) and gene prioritisation (*17*), we ranked these genes based on the mean phenotypic similarity between the three cases who underwent HPO coding and HPO terms derived from the OMIM (Online Mendelian Inheritance in Man) and MGI (Mammalian Genome Informatics) databases. This approach ranked *SRC* at the top of the 31 candidate genes for which term data were available (Fig. 1B). *SRC* encodes the proto-oncogene tyrosine-protein kinase SRC and knock-out of its highly homologous orthologue in mice showed no bleeding and apparently normal platelets (*13*). However, the *Src-/-* mouse presented with increased bone density (osteopetrosis), which is the opposite of the osteoporosis of the pedigree cases (*18*). The proximity of these two related but opposing terms in the ontologies enabled *SRC* to score highest (Fig. 1, B and C). The *SRC* variant had a Combined Annotation-Dependent Depleted (CADD) (*19*) phred score of 34 (Fig. 1D), and was amongst only 24 of the 67 candidates to be unobserved in 61,486 unrelated subjects from the ExAC database and 2,974 further subjects from our in-house collection. Inspection of the results of sequencing of RNA from blood stem and progenitor cells, including MKs (*20*), showed that the *SRC* transcript ranked amongst the highest according to the probability of being over-expressed in MKs compared to the other seven cell types (posterior probability=0.47) (Fig. 1E). Finally, the *SRC* variant c.1579G>A was found by Sanger sequencing to co-segregate with the disease phenotype in three additional cases and was absent from unaffected relative 21 (DNA from the remaining four cases was not available) (p = 0.03125) (fig. S1). The four independent sources of evidence thus established the variant coding glutamic acid (E) 527 lysine (K) in SRC's kinase domain as the primary causative candidate for this novel syndrome.

Platelet phenotypes and functional effect of the E527K *SRC* variant

Morphological examination of platelets from three cases by electron microscopy (EM) shows heterogeneous sized platelets (Fig. 2, A and B and fig. S2) with features of Gray Platelet Syndrome (GPS) as they have a paucity of α-granules (Fig. 2C) and abundant vacuoles. This was confirmed by quantitative blotting showing reduced levels of α-granule-stored von Willebrand factor (VWF) and thrombospondin 1 (TSP1) (Fig. 2D). Patients with classic GPS caused by recessive mutations in *NBEAL2* (*21*) have a bleeding disorder because the cargo of α-granules is essential for maintaining hemostasis. GPS patients and *Nbeal2-/-* mice show premature myelofibrosis leading to extramedullary formation of blood cells and splenomegaly. This fibrosis is likely caused by pro-inflammatory MKs (*22*). However, some other features of our cases are not typical for classic GPS, such as presence of both abnormally large and small platelets without any type of granules and lacking internal membranes (Fig. 2A and fig. S2). As expected, activation with collagen or the more potent rattlesnake venom convulxin causes appearance of the α-granules' membrane protein P-selectin (CD62P) on the platelet outer-membrane. However, in keeping with the EM and blot-ascertained paucity of α-granule proteins, the amounts of P-selectin after activation were significantly reduced (collagen: p = 1.09 x 10^{-7}; convulxin: p = 2.65 x 10^{-3}) (Fig. 2E and fig. S3**)**.

4

The mutated residue E527 is three amino acids upstream from Y530 in the C-terminal tail of the SRC kinase domain (Fig. S4). The negatively charged residue is conserved for SRC across 100 vertebrate species (PhyloP (23) p = 7.48 x 10^{-5}) and also in the remaining seven SFKs. Differential phosphorylation between Y530 and Y419 regulates switching of SRC between inactive and active states. SRC is predominantly in the former state because Y530 mediates an interaction with the SH2 domain, which effectively folds SRC into a closed inaccessible bundle. Autophosphorylation of Y419 enforces an active configuration by displacing it from the kinase cleft, allowing substrates to gain access (24, 25). The three-dimensional structure of SRC reveals, interestingly, that the C-terminal tail (denoted here as the Velcro-strap) carries an extremely high temperature factor, indicating high flexibility (red color in Fig. 2F and fig. S5). SRC is maintained in the inactive state because this Velcro-strap stabilizes the SH2 domain through electrostatic interactions between residues E527 with R159 and of Y530 with the side-chains of R158, R178 and main-chain atoms of T181 and E182 (Fig. 2F, left). Modelling of the consequences of the mutation at residue 527 indicates that the favored interaction with R159 may be lost and is possibly replaced by an alternative electrostatic interaction with the neighboring negatively charged glutamic and aspartic acids at residues 520 or 521, respectively (Fig. 2F, right). This assumed reconfiguration caused by the mutation may favor SRC to adopt a constitutively active state. We challenged this hypothesis by testing the wild-type (WT) and mutant forms with the Phosp3D program (26), which predicted Y530 phosphorylation only for WT SRC. An *in vitro* kinase assay using GST-tagged SRC proteins indeed show a higher kinase activity for E527K compared to WT SRC (p = 2.2 x 10^{-16}) that can partially be inhibited by addition of SRC inhibitor-1 (Fig. 2G, left). Immunoblot analysis of these GST-tagged SRC proteins confirms the differential phosphorylation pattern between WT and E527K SRC using specific Y530 and Y419 antibodies that recognize inactive and active SRC, respectively (Fig. 2G, right).

E527K *SRC* expression in patients' platelets and transfected cells

The predicted consequence of the E527K mutation on SRC's function was explored using SRC activity-specific antibodies. Immunoblot analyses of platelets show increased levels of active SRC in mutation carriers versus controls after accounting for global genotype and antibody effects (p = 9.42 x 10^{-15}) and irrespective of the activation state of the platelets (Fig. 3A and supplementary materials). Platelets from unaffected pedigree member 21 could only be tested under non-activated conditions but show comparable SRC levels to those in control platelets (Fig. 3A). SRC activation in platelets of affected cases is accompanied by increased overall tyrosine phosphorylation irrespective of the platelet activation status (Fig. 3B and fig. S6). These platelet abnormalities are independent of the bone marrow changes, as case 19 has no myelofibrosis. The expression of three other SFK members (*FGR, FYN* and *YES*) was normal in platelets of the cases (fig. S7).

Platelets from cases are assumed to contain both WT and mutant SRC. To determine the activation state of mutant SRC in the absence of WT, we turned to COS-7 cells that do not express SRC protein at detectable levels. We transfected the cells with mutant, WT and a mixture of mutant and WT protein to simulate a homozygous WT, homozygous mutant and

5

heterozygous genotype. Reactivity of the antibody against Y530 was negatively associated with the genotype (p = 2.39 x 10^{-6}) while reactivity of the antibody against Y419 was positively associated with the genotype (p = 0.014) (Fig. 3C, left). In cells harboring the mutant nearly all SRC reacted with the antibody against Y419 while reactivity with the antibody against Y530 was negligible (Fig. 3C, right and supplementary materials), as shown previously for activating Y527F mutant chicken SRC (corresponding to residue 530 in humans) (*27*). Interestingly, co-transfection of WT and mutant protein was sufficient to induce substantially increased Y419 levels (Fig. 3C). Transfections of WT and mutant SRC in the presence of SRC inhibitor-1 resulted in significantly lower Y419 SRC levels for mutant transfected cells (p = 7.78 x 10^{-5}) while no changes were observed for WT transfected cells (Fig. 3D). This is in agreement with the previously described observation that autophosphorylation and dephosphorylation of Y419 are directly correlated with the level of SRC activity (*28*).

Defective megakaryopoiesis due to E527K SRC

Of the eight SFKs (*7*), SRC exhibits the strongest over-expression in MKs relative to other blood cell progenitors (table S1). Detailed examination of several bone marrow biopsy samples from case 35 showed trilineage dysplasia with a high number of MKs having dysplastic features and hypo-lobulated nuclei (table 1), which are hallmarks of immaturity. Thrombopoietin (TPO) is the pivotal growth and differentiation factor for MKs and its receptor MPL is present on blood stem cells, MKs and platelets (*29*). Platelets from *SRC*-mutant cases with myelofibrosis express normal MPL levels (fig. S8) but their plasma contains elevated TPO levels (Fig. 4A). Elevated plasma TPO levels are exceptional and characteristic for *MPL*-inactivating mutations which cause congenital amegakaryocytic thrombocytopenia (CAMT) (*30*). However, hematopoietic stem cells from CAMT cases cannot form MKs in the presence of TPO while *SRC*-mutant blood stem cells produced more MK but also more myeloid (GEMM) colonies (Fig. 4B) than control cells, which is in agreement with the bone marrow biopsy findings. Blood stem cells of three pedigree cases, when cultured with TPO, formed numerous immature MKs with almost no proplatelets (Fig. 4C). Proplatelet formation for unaffected Case 21 was normal (Fig. 4C). Interestingly, addition of the SRC inhibitor-1 starting from culture day 1 could partially restore proplatelet formation for affected Case 31 while it had no discernible effect on control or unaffected cells (Fig. 4C). Immunostaining for active SRC (Y419) of day 12 MKs showed a significantly increased number of positively-stained MKs for the patients compared to control MKs that only showed a very weak perinuclear signal (Fig. 4D). In contrast, staining of E527K MKs showed evidence of active SRC located in rosette-like structures (Fig. 4D). The oncoprotein v-Src is known to alter the actin cytoskeleton and induce focal cell adhesions and podosomes that induced changes in cell morphology and migration (*6, 31*). Staining of fibroblasts from patient 35 indeed also showed active SRC staining near focal adhesions, which is not observed in control cells (fig. S9).

To further ascertain the effect of E527K on megakaryopoiesis, blood stem cells from normal donors were transduced with lentiviral vectors that overexpress GFP, WT and mutant SRC. Immunoblot analysis of differentiated MKs showed high levels of Y530 and Y419 under WT and mutant transduced conditions, respectively (Fig. 4E). As observed in platelets,

increased overall tyrosine phosphorylation was present in mutant MKs while in the WT only a single band was present that likely corresponds to Y530 SRC (Fig. 4E). Ploidy analysis of the transduced MKs showed weak evidence of an increased number of immature (2N) MKs in E527K overexpression conditions compared to WT (Fig. 4F and fig. S10). Quantification of proplatelet formation showed significantly reduced proplatelet-forming MKs for E527K transduced conditions compared to WT (Fig. 4G and fig S11). Immunostaining for total or Y419 SRC was weak for non-transduced cells (fig. S12) while strong staining was present for the SRC-transduced conditions. Mutant-transduced MKs recapitulated the presence of rosette-like structures that partially co-stain with filamentous actin, which were absent from WT (Fig. 4H and fig. S12). In general, E527K-transduced MKs have a strongly altered actin organisation even when they have a tendency to form proplatelets (fig. S12C). Cortactin (or cortical-actin binding protein) is normally located in the cytoplasm but upon tyrosine-phosphorylated induced activation, it locates to podosomes where it promotes actin reorganisation (*32*). Staining of the lentiviral transduced MKs for cortactin showed similar types of rosette structures in E527K-transduced MKs while WT-transduced cells have significantly reduced numbers of cortactin-positive podosomes for different time points during MK differentiation (fig. S13).

Blood and bone phenotypes in Src-E525K overexpression zebrafish
We injected zebrafish *src* mRNA containing the E525K mutation (corresponding to E527K in humans) or the WT sequence to study the formation of blood cells and bone. Overexpression of mutant *src* resulted in abnormal early primitive hematopoiesis between 16 and 24 hours post fertilization (hpf) with a migration defect of GATA1-positive stem cells (fig. S14). Mutant embryos at 3 days post fertilization (dpf) have normal formation of red cells (fig. S14) while thrombocyte numbers appear to be reduced (Fig. 5A). Counting of thrombocytes by flow cytometry showed 493 versus 852 CD41$^+$ thrombocytes in 10 pooled mutant *src* and control embryos respectively. Immunoblot analysis of 15 embryo lysates showed reduced levels of GFP-positive thrombocytes though quantification analysis of a triplicate injection experiment showed only a trend for lower GFP levels for E525K mutant embryos that did not reach statistical significance (Fig 5A). SRC levels did not differ between buffer (control) or RNA-injected embryos (Fig. 5A), which probably indicates that the RNA overexpression is not detectable at 3 dpf. E525K *src*-injected embryos in the presence of SRC inhibitor-1 had lower tyrosine phosphorylation levels and higher GFP levels compared to embryos without the inhibitor. Interestingly, E525K *src* -injected embryos at 5 dpf have significantly smaller bones compared to WT-injected embryos and this change was absent when the SRC inhibitor-1 was added to the fish water (Fig. 5B). Splice morpholino-induced *src*-depleted embryos show normal thrombocyte formation but have larger bones (fig. S15).

DISCUSSION
Inherited thrombocytopenia (IT) with early onset myelofibrosis was the clinical diagnosis for cases of the presented pedigree upon referral for genetic studies. ITs are a heterogeneous group of diseases caused by at least 20 different genes that cause alterations in transcription regulation, TPO signalling, cytoskeletal organization, granule trafficking and receptor signalling

in patients' MKs (*33*), which result in lower platelet counts with either larger, smaller or normal sized platelets. None of the previously described ITs matched entirely with the platelet phenotype in our cases. They were characterised by presence of both unusually small and large platelets, of which some have a GPS-like phenotype, elevated plasma TPO levels, increased numbers of MKs in the marrow and differentiated *in vitro* trilineage dysplasia and myelofibrosis. All known genetic causes for ITs were excluded (*34*). A detailed clinical description using HPO terms (*34*) that also included non-hematological terms to describe the mild bone defects in cases of this pedigree, such as decreased bone mineral density, fractures and tooth loss, assisted in the unique gene discovery approach used to discover the cause of this new syndrome. Our discovery has been made possible by comparison of HPO terms of the cases with Mouse Phenotype Ontology-mapped HPO terms from 7,541 strains. Though Src-deficient mice have normal platelets, they were described as manifesting with 8 terms that overlapped with the HPO terms for the human cases, including abnormality of teeth and abnormal bone mineral density (*13*). In addition, genome data analysis was empowered by making use of a large cohort of patients with inherited bleeding and platelet disorders (*34*). The selection of the *SRC* gene variant E527K as the most likely candidate was also aided by using the Blueprint blood cell progenitor RNA-seq dataset (*20*).

The E527K variant was predicted by modeling and shown by *in vitro* studies to result in a constitutively active kinase that causes increased overall tyrosine phosphorylation levels in platelets and megakaryocytes. As SRC was not previously expected to play a role in megakaryopoiesis and platelet formation, we performed extensive studies using patient-derived and E527K lentiviral transduced stem cells. These data show that active E527K SRC results in more immature MKs and a defect in proplatelet formation with extensive alterations in the actin cytoskeleton and podosome structures when adhered to fibrinogen. As found for constitutively active v-Src-transformed cells, these actin-based dynamic protrusions of the plasma membrane were typically present at sites where the cells adhere to the extracellular matrix (ECM) (*5*, *6*) but they also mediate cell migration (*35*). Podosomes have since been studied as dynamic F-actin-based adhesion spots associated with motile cells of the myeloid lineage such as macrophages (*36*), dendritic cells (*37*) and megakaryocytes (*38*). These studies were all conducted using blood cells from Wiskott-Aldrich syndrome (WAS) patients and showed reduced podosome formation and cell motility. WAS is an X-linked disorder characterized by eczema, microthrombocytopenia and severe immunodeficiency due to *WAS* gene mutations. Sabri et al (*38*) found that *Was* knockout mice have MKs that almost completely lack actin-rich podosomes after interaction with the bone matrix component collagen I, which results in enhanced proplatelet formation and MKs that shed platelets within the bone marrow space (*38*). *WAS* codes for a protein that regulates the actin cytoskeleton via the Arp2/3 complex. Interestingly, WAS is known to be tyrosine phosphorylated though it is unclear whether this actually influences WAS activities (*39*). It was recently shown by *in vitro* studies that podosomes in normal MKs can degrade the ECM and are therefore predicted to have a pivotal role in MK motility and ECM remodelling, including activities such as extension of proplatelet arms across the membrane of a sinusoidal vessel (*40*). We found no evidence of proplatelet formation within the bone marrow of our cases and observe the opposite phenotype of increased podosome

8

formation and reduced proplatelet formation, which might indicate that the exact number or life time of podosomes might be critical for proplatelet formation. Podosomes have also been described in osteoclasts (*41*) and have been shown to precede bone resorption (*42*). The opposite bone phenotypes of osteopetrosis in Src-deficient mice and osteoporosis in our cases might also be due to the critical balance of podosome numbers.

Finally, we have tried to mimic this novel disorder in a zebrafish model. Though we clearly observed the opposite bone defect in Src-depleted versus E525K-Src-injected embryos with significantly larger and smaller bones, respectively, the blood cell phenotype is less obvious. However, this might be due to the fact that though the molecular control of primitive hematopoiesis is conserved between humans and zebrafish (*43*), several aspects known for human megakaryopoiesis including MK endomitosis and proplatelet formation with loss of the nucleus, are not present in zebrafish (*44*). At 3 dpf, embryos develop high and low GFP (CD41+) positive cells in the CHT tail region that consists of immature and mature thrombocytes with a nucleus, respectively.

Another study limitation in addition to these important differences between human and zebrafish megakaryopoiesis is the fact that we have not yet identified members of other pedigrees with a similar genetic defect. However, we have shown that the profound proplatelet formation defect found in MKs from the affected cases can be rescued using an SRC inhibitor and this observation can be replicated in normal stem cells transduced with the E527K lentiviral vector. The definitive experiment would be to rescue the proplatelet defect in MKs differentiated from a patient-derived iPS line via genetic correction of E527K.

In conclusion, our genome sequencing analysis approach based on detailed HPO coding, progenitor cell expression data, pathogenicity predictions and co-segregation studies was used to discover a germline gain-of-function variant of *SRC* linking the first ever discovered oncogene, which has been studied for decades, with a new syndrome.

MATERIALS AND METHODS

Study design

The overall objective of this study was to identify and characterise the genetic defect causing thrombocytopenia, bleeding and myelofibrosis in affected cases of the presented pedigree. A variant encoding E527K in SRC arose as a primary candidate. Subsequent studies were performed to identify the defective pathway, analyse the platelet and MK defect using patients' cells, replicate the phenotypes using lentiviral transduced stem cells and a zebrafish model and apply a rescue strategy using SRC inhibitor-1. Platelet and MK defects were studied in patients with (cases 31 and 35) and without (case 19) signs of myelofibrosis. The following experiments were performed blinded: quantification of proplatelet formation using lentiviral transduced stem cells and the quantification of bone length and area for all zebrafish experiments. All data are included (no outlier values were excluded).

HPO node abbreviations

Abdom. Orgs.: Abdominal organs; AG: Abnormal α-granules; AGD: Abnormal α-granule distribution; BBFT: Abnormality of blood and blood-forming tissues; BMT: Bleeding with minor or

no trauma; Endo.: Endocrine; Gen.: Genitourinary; IMPV: Increased mean platelet volume; MCLBM: Abnormality of multiple cell lineages in the bone marrow; Min.: Mineral; Morph.: Morphology; NAG: Abnormal number of α-granules; OCS: Abnormal surface-connected open canalicular system; PA: Phenotypic abnormality; Plt(s): Platelet(s); Recur. Frac.: Recurrent fractures; Skel.: Skeletal; Subcut. Hem.: Subcutaneous hemorrhage; Sys.: System; TCP: Thrombocytopenia. Certain term definitions have been shortened for conciseness (e.g. "Abnormality of the teeth" is shown as "Teeth").

Case information and informed consent

Clinical information of the pedigree is reported in **Fig. 1A** and **Table 1**. Pedigree members signed informed consent to participate in the BRIDGE-BPD study and the enhanced clinical and laboratory phenotyping studies. The Ethics Committee of the University Hospital Leuven approved the study (ref. ML3580). Cases enrolled to the BRIDGE-BPD (UK REC10/H0304/66) and to the NIHR BioResource-Rare Diseases (UK REC 13/EE/0325)/BPD studies after providing informed written consent.

DNA extraction, sequencing and variant calling

Genomic DNA was isolated from venous blood or saliva obtained from cases at enrolment or retrieved from the sample archive. Extracted DNA was quality controlled by gel electrophoresis and by three independent measurements of DNA concentration, namely Picogreen (Life Technologies Ltd), Qubit (Life Technologies) and Glomax (Promega). Whole exome sequencing (WES) for DNA samples in BRIDGE was performed as described previously (*15*). Whole genome sequencing (WGS) libraries were prepared using a TruSeq DNA PCR-free protocol and sequenced in 125 bp paired-end reads by Illumina Cambridge Inc. such that 95% of the GRCh37 reference genome had a read coverage of at least 15X. The G to A variant at position 36,031,750 of chromosome 20 was confirmed by Sanger sequencing (StarSEQ) using PCR-amplified fragments (F-cccactttcctcaccggagcc and R-ctcgcccctggcaattcagccc).

Variant annotation and filtering

A multi-sample VCF was annotated with allele frequencies from ExAC release 0.2 containing genotypes for up to 61,486 samples. CADD scores and functional impact with respect to Ensembl 70 were obtained using snpEff (*45*). Variants with a frequency greater than 0.001 in ExAC were removed by filtering. The allele frequency threshold for removing variants in our in-house collection of 2,974 subjects was set to 0.01.

Literature-based gene prioritisation

In order to obtain the ranking by phenotypic relevance of the 67 candidate variants, we constructed HPO profiles for each of the corresponding genes and computed the mean phenotypic similarity (*46*) of cases 13, 31 and 35 to those profiles. The phenotypic similarity used information content assigned to each term derived from 856 unrelated individuals with inherited BPDs of unknown molecular etiology in the BRIDGE study. The profiles were constructed by combining HPO terms associated with each gene (whereby they were assigned

10

terms associated with diseases (*46*) linked to the gene through OMIM) and HPO terms derived from mapping MPO terms associated with corresponding mouse models (*47*) through the cross-species ontology Uberpheno (*48*). Terms from OMIM and/or MPO were available for 31 of the 67 candidate genes.

Identification of genes specifically overexpressed in megakaryocytes
We used the MMSEQ (*49*) expression estimates from the Blueprint Consortium progenitors dataset (*20*) to perform a comparison of nine different models using MMDIFF (*50*). We compared a baseline model with single mean across all cell types, to which we assigned a prior probability of 0.5, with eight alternative models in which the mean for one cell type was assumed to differ from the mean for all other cell types combined. We assigned a prior probability of 1/16 to each of the alternative specific models. The probability of the model postulating a megakaryocyte (MK)-specific mean expression level was multiplied by the sign of the fold change between the MK samples and the other samples in order to rank genes by their probability of being overexpressed specifically in MKs.

Probability of segregation between genotype and phenotype
Assuming that a single founder introduced the rare variant into the pedigree, the probability of the genotypes (obtained by Sanger sequencing) of additional members of the pedigree given the pedigree structure and the initial genotyping results (called by high-throughput sequencing) under the null hypothesis that the variant is not in linkage with the phenotype is given by

$$P(G|Z) = \sum_{F} P(G \mid F_i = 1, Z)P(F_i = 1|Z),$$

where the random variable $G_i = 1$ if subject i has a copy of the mutant allele as verified by Sanger sequencing and *0* otherwise; $Z_i = 1$ if subject i has a copy of the mutant allele as verified initially by WES/WGS and *0* otherwise; and $F_i = 1$ if and only if subject i is a founder and introduced the variant into the pedigree. We derive the probability of the co-segregation results in the pedigree under the null of no linkage as $2 \cdot (\frac{1}{2})^5(\frac{1}{2}) = 0.03125$, given $G_i = 1$ if $i \in$ {*19,23,25*}, $G_i = 0$ if $i = 21$ and G_i is unknown otherwise; $Z_i = 1$ if $i \in$ {*13,31*} and $Z_i = 0$ otherwise.

Reagents
The following antibodies were used: rabbit antibodies against total (pan), Y419 and Y530 forms of SRC (all from Life Technologies), rabbit antibody SRC that cross-reacts with Zebrafish (2108, Cell Signaling), tyrosine phosphorylation antibody 4G10 (Millipore), rabbit antibody for Cortactin (H222, Cell Signaling), phalloidin-rhodamine for F-actin (Sigma) and mouse monoclonal against GAPDH (clone 4G5; Fitzgerald Industries International). SRC inhibitor-1 (Sigma) was used at 500 nM final concentration for all experiments except that 2 µM was added to zebrafish water.

Functional and morphological platelet studies

EDTA anticoagulated blood was analyzed on an automated full blood cell analyser to determine blood cell counts and other red cell indices. Platelet-rich plasma (PRP) was prepared by centrifugation (15 minutes at 150g) of whole blood anticoagulated with 3.8% trisodium citrate (9:1). PRP was used for functional studies and EM, as described (*51*). Briefly, aggregation studies were carried out by adding Horm collagen, ristocetin, TRAP6, arachidonic acid, U46619 and ADP at the indicated concentrations (**Table 1**). ATP secretion tests were performed after stimulation of platelets with Horm collagen. The expression of P-selectin (CD62P) on the platelet outer-membrane was measured by flow cytometry as described under basal and activated conditions with Horm collagen and convulxin (*52*). We used the Cell Diva software for two-color immunofluorescence acquisition on a FACS Canto II flow cytometer (BD Biosciences) and FlowJo software for analysis and presentation (Tree Star Inc). See below for details of the statistical analysis.

Western blot of total platelet lysates for α-granule markers
Protein lysates were obtained from platelets as described previously (*51*). Protein fractions were resolved by SDS/PAGE and blots were incubated with the following rabbit polyclonal antibodies: homemade anti-Tsp1 (*53*), anti-Vwf-HRP (Dako) and anti-integrin β3 (Itgb3) antibody (H-96) (Santa Cruz Biotechnology). Membranes were next incubated with HRP-conjugated secondary antibody and staining was performed with the ECL detection reagent (Life Technologies). Chemiluminescent blots were imaged with the ChemiDoc MP imager and the ImageLab software version 4.1 (Bio-Rad) was used for image acquisition. Unaltered blots are included in fig. S16-20.

Western blot of basal and collagen-activated washed platelets for SRC quantification
To obtain washed platelets, PRP from citrate anticoagulated blood was diluted with 2 volumes of acid-citrate-dextrose (ACD) supplemented with prostaglandin E1 (Pfizer) at final concentration of 1 µg/ml. After centrifugation, platelets were resuspended in Tyrode-HEPES buffer (10 mM HEPES [pH 7.5], 12 mM NaHCO3, 137 mM NaCl, 2.7 mM KCl, 5 mM glucose, 1 MgCl) and the concentration was adjusted to 250×10^3 platelet/µL and supplemented with a pharmaceutical blocking compound of $\alpha_{IIb}\beta_3$ (Itga2b/Itgb3) (1/100 of Integrilin, Millennium Pharmaceuticals). Washed platelets (300 µL) were activated with 2 µg/ml Horm collagen or 10 µM ADP for different time intervals without stirring, centrifuged and the pellet was dissolved in 100 µL SDS-phospho loading buffer supplemented with protease inhibitors, 2mM NaF and 2mM NaVO$_3$. Gels were loaded with 5 µL for each condition. Chemiluminescent blots were imaged with the ChemiDoc MP imager (Bio-Rad) and the ImageLab software version 4.1 was used for image acquisition and densitometric analysis. Statistical analysis is described in the supplementary materials. Unaltered blots are included in fig. S16-20.

SRC kinase assay
The complete coding region of wild-type (E527) and mutant (527K) SRC was cloned as a GST-tagged construct in the pGEX-4T-2 vector (GE Healthcare Biosciences, Pittsburgh, PA, USA). Purified GST-SRC proteins using glutathione sepharose beads were used in combination with

the HTScan SRC kinase assay kit (Cell Signaling Technology) using the colorimetric ELISA approach as described by the manufacturer (*54*).

Cloning, COS-7 transfections and Western blot analysis
Wild-type and mutant SRC was cloned as a Myc-tagged construct in the pSecTag2 vector (Life Technologies). COS-7 cells were transfected with jetPRIME DNA and siRNA transfection reagent according to the manufacturer's instructions (Polyplus Transfection). Cells were lysed 48 hours after transfection in SDS-phospho buffer and SRC and GAPDH blots were done as described for washed platelets. Transfections and all blots were done in triplicate using 2µg E527, 527K or empty vector (V) or a combination of 1µg E527 with 1µg 527K. Unaltered blots are included in fig. S16-20.

Thrombopoietin measurement
TPO levels were determined in citrated plasma using a commercially available kit (R & D Systems Inc.).

Colony assays
CD34$^+$ HSCs were isolated from peripheral blood from case 31 and an unrelated healthy control by magnetic cell sorting (Miltenyi Biotec) and colonies were grown as described (*55*, *56*). Briefly, CD34$^+$ cells (1 x 10^4 and 1 x 10^3; respectively for MK and GEMM) were cultured in triplicate in Megacult-C 04973 (MK) and Methocult 04964 (GEMM), according to the manufacturer's instructions (StemCell Technologies). The total number of MK and GEMM colonies were counted 7 and 12 days later using a light microscope (Leica DM RBE). Colonies were blindly counted.

Suspension megakaryocyte cultures, lentiviral transductions, proplatelet and ploidy analysis
CD34$^+$ HSCs were isolated from peripheral blood from pedigree cases 19, 21, 31 and 35 and three unrelated healthy controls as per above. WT and E527K SRC were cloned to replace the eGFP-T2A-fLuc reporter cassette in pCH-SFFV-eGFP-P2A-fluc transfer plasmid (*57*). The pCH-SFFV-eGFP-P2A-fluc was used as control. Vector particles were produced and concentrated by the Leuven Viral Vector Core as described before (*57*). About 1x10e6 CD34+ cells isolated from unrelated normal blood buffy coats were transduced for 48 hours with these vectors and differentiated as described (*56*) to MKs. About 1 x 10^6 CD34+ cells isolated from unrelated normal blood buffy coats were transduced for 48 hours with lentiviral vectors in the presence or absence of SRC inhibitor. CD34$^+$ cells were cultured in StemSpan SFEM medium with or without SRC inhibitor (Stemcell Technologies), supplemented with 20 ng/ml TPO, 10 ng/ml Scf, 10 ng/ml Il-6 and 10 ng/ml Flt-3 (Peprotech) as described (*56*). At day 12 of culture, proplatelet forming MKs were counted and, for immunostaining experiments, MKs were placed overnight on fibrinogen-coated coverslips at different days of culture (8, 12 and 13). Stainings were photographed at 63x magnification with a Zeiss Axiovert microscope and captured with Zeiss Axiovision. Ploidy analysis was performed at day 12 by flow cytometry as described (*56*).

Functional genetics in zebrafish

Tg(cd41:EGF) (*44*) (gift from Dr L. Zon (Hematology Division, Brigham and Women's Hospital's, Boston) zebrafish embryos were injected at the 1-cell stage with a *src* splice morpholino (tgtagtctgctttacctaatatgcc), developed by Gene-Tools LLC (Philomath), or wild type or E525K mutant src mRNA (ENSDARG00000008107). The analysis of thrombocyte, erythrocyte and GATA1-positive hematopoietic stem cell formation was analysed as we described before (*55*, *58*). Cartilage structures were stained at 5 dpf with alcian blue as described (*59*) and quantification of bones was performed as described (*59*). All animal protocols were approved by the Ethical Committee of the Katholieke Universiteit Leuven.

Fibroblast staining

Skin fibroblasts were grown in DMEM/F12 and passage 4 were used for immunostaining. Adherent fibroblasts were washed and fixed with 4% paraformaldehyde in cytoskeleton buffer (0.1M PIPES, 2M glycerol, 1mM EDTA, 1mM MgCl2, pH 6.9), and permeabilized for 15 minutes with 0.2\% triton X-100 (Roche) at room temperature. Images were performed on a Zeiss Axiovert 100M confocal microscope (Carl Zeiss inc.).

Statistical analysis of flow cytometry and Western blot data

Analysis of flow cytometric measurements of P-selectin (CD62P) on the outer-membrane of platelets, OD_{450} readout of the ELISA assay on bacterial cells and Western blot data from platelet lysates and COS-7 cells was done by fitting the linear mixed models described in Supplementary Materials. Regression coefficients were estimated using the lme4 R package (*60*) and associated *p*-values were computed using likelihood ratio tests. Statistical analysis for colony assays, proplatelet formation, quantification of Y419+ MKs, MK ploidy and zebrafish data was performed using Graphpad Prism 6. Details on sample sizes, statistical methods and *p*-values are listed in table S2.

SUPPLEMENTARY MATERIALS

Supplementary Materials and Methods
- Statistical analysis of CD62P FACS data
- Statistical analysis of OD_{450} ELISA data
- Statistical analysis of platelet Western blot data
- Statistical analysis of COS-7 transfection Western blot data
- Statistical analysis of COS-7 transfection Western blot data with inhibitor

Fig. S1. Sanger sequencing chromatograms
Fig. S2. Electron microscopy images
Fig. S3. Representative histograms of CD62P release from α-granule membranes
Fig. S4. Amino acid conservation across species and human SFKs
Fig. S5. Protein structure of SRC
Fig. S6. Antiphosphotyrosine immunoblots
Fig. S7. Immunoblot analysis of FGR, YES and FYN

REFERENCES

1. P. Rous, A TRANSMISSIBLE AVIAN NEOPLASM. (SARCOMA OF THE COMMON FOWL.). *J. Exp. Med.* **12**, 696–705 (1910).

2. P. Rous, A SARCOMA OF THE FOWL TRANSMISSIBLE BY AN AGENT SEPARABLE FROM THE TUMOR CELLS. *J. Exp. Med.* **13**, 397–411 (1911).

3. T. Hunter, Discovering the first tyrosine kinase. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 7877–82 (2015).

4. W. T. Chen, Proteolytic activity of specialized surface protrusions formed at rosette contact sites of transformed cells. *J. Exp. Zool.* **251**, 167–85 (1989).

5. D. A. Murphy, S. A. Courtneidge, The "ins" and "outs" of podosomes and invadopodia: characteristics, formation and function. *Nat. Rev. Mol. Cell Biol.* **12**, 413–26 (2011).

6. R. B. Irby, W. Mao, D. Coppola, J. Kang, J. M. Loubeau, W. Trudeau, R. Karl, D. J. Fujita, R. Jove, T. J. Yeatman, Activating SRC mutation in a subset of advanced human colon cancers. *Nat. Genet.* **21**, 187–90 (1999).

7. T. J. Yeatman, A renaissance for SRC. *Nat. Rev. Cancer.* **4**, 470–80 (2004).

8. G. S. Martin, The hunting of the Src. *Nat. Rev. Mol. Cell Biol.* **2**, 467–75 (2001).

9. M. Breccia, G. Alimena, Second-Generation Tyrosine Kinase Inhibitors (Tki) as Salvage Therapy for Resistant or Intolerant Patients to Prior TKIs. *Mediterr. J. Hematol. Infect. Dis.* **6**, e2014003 (2014).

10. H. M. Kantarjian, N. P. Shah, J. E. Cortes, M. Baccarani, M. B. Agar- wal, M. S. Undurraga, J. Wang, J. J. K. Ipiña, D.-W. Kim, M. Ogura, C. Pavlovsky, C. Junghanss, J. H. Milone, F. E. Nicolini, T. Robak, J. Van Droogenbroeck, E. Vellenga, M. B. Bradley-Garelik, C. Zhu, A. Hochhaus, Dasatinib or imatinib in newly diagnosed chronic-phase chronic myeloid leukemia: 2-year follow-up from a randomized phase 3 trial (DASISION). *Blood.* **119**, 1123–9 (2012).

11. A. Golden, S. P. Nemeth, J. S. Brugge, Blood platelets express high levels of the pp60c-src-specific tyrosine kinase activity. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 852–6 (1986).

12. J. B. Levy, T. Dorai, L. H. Wang, J. S. Brugge, The structurally distinct form of pp60c-src detected in neuronal cells is encoded by a unique c-src mRNA. *Mol. Cell. Biol.* **7**, 4142–5

(1987).

13. P. Soriano, C. Montgomery, R. Geske, A. Bradley, Targeted disruption of the c-src proto-oncogene leads to osteopetrosis in mice. *Cell*. **64**, 693–702 (1991).

14. Y. A. Senis, A. Mazharian, J. Mori, Src family kinases: at the forefront of platelet activation. *Blood*. **124**, 2013–24 (2014).

15. S. K. Westbury, E. Turro, D. Greene, C. Lentaigne, A. M. Kelly, T. K. Bariana, I. Simeoni, X. Pillois, A. Attwood, S. Austin, S. B. Jansen, T. Bakchoul, A. Crisp-Hihn, W. N. Erber, R. Favier, N. Foad, M. Gattens, J. D. Jolley, R. Liesner, S. Meacham, C. M. Millar, A. T. Nurden, K. Peerlinck, D. J. Perry, P. Poudel, S. Schulman, H. Schulze, J. C. Stephens, B. Furie, P. N. Robinson, C. van Geet, A. Rendon, K. Gomez, M. A. Laffan, M. P. Lambert, P. Nurden, W. H. Ouwehand, S. Richardson, A. D. Mumford, K. Freson, BRIDGE-BPD Consortium, Human phenotype ontology annotation and cluster analysis to unravel genetic defects in 707 cases with unexplained bleeding and platelet disorders, *Genome medicine*. **7**, 36 (2015).

16. M. H. Schulz, P. Krawitz, S. Bauer, S. Do, S. Ko, C. E. Ott, C. Mundlos, D. Horn, S. Mundlos, P. N. Robinson, S. Köhler, S. Dölken, Clinical diagnostics in human genetics with semantic similarity searches in ontologies. *Am. J. Hum. Genet.* **85**, 457–464 (2009).

17. P. N. Robinson, S. Köhler, A. Oellrich, K. Wang, C. J. Mungall, S. E. Lewis, N. Washington, S. Bauer, D. Seelow, P. Krawitz, C. Gilissen, M. Haendel, D. Smedley, Improved exome prioritization of disease genes through cross-species phenotype comparison. *Genome Res.* **24**, 340–8 (2014).

18. F. Lazner, M. Gowen, D. Pavasovic, I. Kola, Osteopetrosis and osteoporosis: two sides of the same coin. *Hum. Mol. Genet.* **8**, 1839–46 (1999).

19. M. Kircher, D. M. Witten, P. Jain, B. J. O'Roak, G. M. Cooper, J. Shendure, A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **46**, 310–5 (2014).

20. L. Chen, M. Kostadima, J. H. A. Martens, G. Canu, S. P. Garcia, E. Turro, K. Downes, I. C. Macaulay, E. Bielczyk-Maczynska, S. Coe, S. Farrow, P. Poudel, F. Burden, S. B. G. Jansen, W. J. Astle, A. Attwood, T. Bariana, B. de Bono, A. Breschi, J. C. Chambers, F. A. Choudry, L. Clarke, P. Coupland, M. van der Ent, W. N. Erber, J. H. Jansen, R. Favier, M. E. Fenech, N. Foad, K. Freson, C. van Geet, K. Gomez, R. Guigo, D. Hampshire, A. M. Kelly, H. H. D. Kerstens, J. S. Kooner, M. Laffan, C. Lentaigne, C. Labalette, T. Martin, S. Meacham, A. Mumford, S. Nürnberg, E. Palumbo, B. A. van der Reijden, D. Richardson, S. J. Sammut, G. Slodkowicz, A. U. Tamuri, L. Vasquez, K. Voss, S. Watt, S. Westbury, P. Flicek, R. Loos, N. Goldman, P. Bertone, R. J. Read, S. Richardson, A. Cvejic, N. Soranzo, W. H. Ouwehand, H. G. Stunnenberg, M. Frontini, A. Rendon, Transcriptional diversity during lineage commitment of human blood progenitors. *Science*. **345**, 1251033 (2014).

21. C. A. Albers, A. Cvejic, R. Favier, E. E. Bouwmans, M.-C. Alessi, P. Bertone, G. Jordan, R. N. W. Kettleborough, G. Kiddle, M. Kostadima, R. J. Read, B. Sipos, S. Sivapalaratnam, P. A. Smethurst, J. Stephens, K. Voss, A. Nurden, A. Rendon, P. Nurden, W. H. Ouwehand, Exome sequencing identifies NBEAL2 as the causative gene for gray platelet syndrome. *Nat. Genet.* **43**, 735–7 (2011).

22. J. A. Guerrero, C. Bennett, L. van der Weyden, H. McKinney, M. Chin, P. Nurden, Z. McIntyre, E. L. Cambridge, J. Estabel, H. Wardle-Jones, A. O. Speak, W. N. Erber, A. Rendon, W. H. Ouwehand, C. Ghevaert, Gray platelet syndrome: proinflammatory megakaryocytes and α-granule loss cause myelofibrosis and confer metastasis resistance in mice. *Blood*. **124**, 3624–35 (2014).

23. G. M. Cooper, E. A. Stone, G. Asimenos, E. D. Green, S. Batzoglou, A. Sidow, Distribution and intensity of constraint in mammalian genomic sequence. *Genome Res.* **15**, 901–13 (2005).

24. S. W. Cowan-Jacob, G. Fendrich, P. W. Manley, W. Jahnke, D. Fabbro, J. Liebetanz, T. Meyer, The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation. *Structure.* **13**, 861–71 (2005).

25. M. K. Ayrapetov, Y.-H. Wang, X. Lin, X. Gu, K. Parang, G. Sun, Conformational basis for SH2-Tyr(P)527 binding in Src inactivation. *J. Biol. Chem.* **281**, 23776–84 (2006).

26. P. Durek, C. Schudoma, W. Weckwerth, J. Selbig, D. Walther, Detection and characterization of 3D-signature phosphorylation site motifs and their contribution towards improved phosphorylation site prediction in proteins. *BMC Bioinformatics.* **10**, 117 (2009).

27. P. Patwardhan, M. D. Resh, Myristoylation and membrane binding regulate c-Src stability and kinase activity. *Mol. Cell. Biol.* **30**, 4094–107 (2010).

28. R. J. Boerner, D. B. Kassel, S. C. Barker, B. Ellis, P. DeLacy, W. B. Knight, Correlation of the phosphorylation states of pp60c-src with tyrosine kinase activity: the intramolecular pY530-SH2 complex retains significant activity if Y419 is phosphorylated. *Biochemistry.* **35**, 9519–25 (1996).

29. F. Wendling, E. Maraskovsky, N. Debili, C. Florindo, M. Teepe, M. Titeux, N. Methia, J. Breton-Gorius, D. Cosman, W. Vainchenker, cMpl ligand is a humoral regulator of megakaryocytopoiesis. *Nature.* **369**, 571–4 (1994).

30. M. Ballmaier, M. Germeshausen, H. Schulze, K. Cherkaoui, S. Lang, A. Gaudig, S. Krukemeier, M. Eilers, G. Strauss, K. Welte, c-mpl mutations are the cause of congenital amegakaryocytic thrombocytopenia. *Blood.* **97**, 139–46 (2001).

31. M. C. Frame, V. J. Fincham, N. O. Carragher, J. a Wyke, v-Src's hold over actin and cell adhesions. *Nat. Rev. Mol. Cell Biol.* **3**, 233–245 (2002).

32. L. C. Kelley, K. E. Hayes, A. G. Ammer, K. H. Martin, S. A. Weed, Revisiting the ERK/Src cortactin switch. *Commun. Integr. Biol.* **4**, 205–7 (2011).

33. K. Freson, A. Wijgaerts, C. van Geet, Update on the causes of platelet disorders and functional consequences. *Int. J. Lab. Hematol.* **36**, 313–25 (2014).

34. S. K. Westbury, E. Turro, D. Greene, C. Lentaigne, A. M. Kelly, T. K. Bariana, I. Simeoni, X. Pillois, A. Attwood, S. Austin, S. B. Jansen, T. Bakchoul, A. Crisp-Hihn, W. N. Erber, R. Favier, N. Foad, M. Gattens, J. D. Jolley, R. Liesner, S. Meacham, C. M. Millar, A. T. Nurden, K. Peerlinck, D. J. Perry, P. Poudel, S. Schulman, H. Schulze, J. C. Stephens, B. Furie, P. N. Robinson, C. van Geet, A. Rendon, K. Gomez, M. A. Laffan, M. P. Lambert, P. Nurden, W. H. Ouwehand, S. Richardson, A. D. Mumford, K. Freson, Human phenotype ontology annotation and cluster analysis to unravel genetic defects in 707 cases with unexplained bleeding and platelet disorders. *Genome Med.* **7**, 36 (2015).

35. H. Schachtner, S. D. J. Calaminus, S. G. Thomas, L. M. Machesky, Podosomes in adhesion, migration, mechanosensing and matrix remodeling. *Cytoskeleton (Hoboken).* **70**, 572–89 (2013).

36. S. Linder, H. Higgs, K. Hüfner, K. Schwarz, U. Pannicke, M. Aepfelbacher, The polarization defect of Wiskott-Aldrich syndrome macrophages is linked to dislocalization of the Arp2/3 complex. *J. Immunol.* **165**, 221–5 (2000).

37. A. Olivier, L. Jeanson-Leh, G. Bouma, D. Compagno, J. Blondeau, K. Seye, S. Charrier, S. Burns, A. J. Thrasher, O. Danos, W. Vainchenker, A. Galy, A partial down-regulation of WASP is sufficient to inhibit podosome formation in dendritic cells. *Mol. Ther.* **13**, 729–37 (2006).

38. S. Sabri, A. Foudi, S. Boukour, B. Franc, S. Charrier, M. Jandrot-Perrus, R. W. Farndale, A. Jalil, M. P. Blundell, E. M. Cramer, F. Louache, N. Debili, A. J. Thrasher, W. Vainchenker, Deficiency in the Wiskott-Aldrich protein induces premature proplatelet formation and platelet production in the bone marrow compartment. *Blood.* **108**, 134–40 (2006).

39. A. Dovas, D. Cox, Regulation of WASp by phosphorylation: Activation or other functions? *Commun. Integr. Biol.* **3**, 101–5 (2010).

40. H. Schachtner, S. D. J. Calaminus, A. Sinclair, J. Monypenny, M. P. Blundell, C. Leon, T. L. Holyoake, A. J. Thrasher, A. M. Michie, M. Vukovic, C. Gachet, G. E. Jones, S. G. Thomas, S. P. Watson, L. M. Machesky, Megakaryocytes assemble podosomes that degrade matrix and protrude through basement membrane. *Blood.* **121**, 2542–52 (2013).

41. F. Saltel, A. Chabadel, E. Bonnelye, P. Jurdic, Actin cytoskeletal organisation in osteoclasts: a model to decipher transmigration and matrix degradation. *Eur. J. Cell Biol.* **87**, 459–68 (2008).

42. A. Zambonin-Zallone, A. Teti, A. Carano, P. C. Marchisio, The distribution of podosomes in osteoclasts cultured on bone laminae: effect of retinol. *J. Bone Miner. Res.* **3**, 517–23 (1988).

43. A. J. Davidson, L. I. Zon, The "definitive" (and "primitive") guide to zebrafish hematopoiesis. *Oncogene.* **23**, 7233–46 (2004).

44. H.-F. Lin, D. Traver, H. Zhu, K. Dooley, B. H. Paw, L. I. Zon, R. I. Handin, Analysis of thrombocyte development in CD41-GFP transgenic zebrafish. *Blood.* **106**, 3803–10 (2005).

45. P. Cingolani, A. Platts, L. L. Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, D. M. Ruden, A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin).* **6**, 80–92.

46. S. Bauer, D. Seelow, D. Horn, P. N. Robinson, S. Ko, S. Mundlos, S. Köhler, The Human Phenotype Ontology: a tool for annotating and analyzing human hereditary disease. *Am. J. Hum. Genet.* **83**, 610–5 (2008).

47. J. A. Blake, C. J. Bult, J. T. Eppig, J. A. Kadin, J. E. Richardson, The Mouse Genome Database: integration of and access to knowledge about the laboratory mouse. *Nucleic Acids Res.* **42**, D810–7 (2014).

48. S. Köhler, S. C. Doelken, B. J. Ruef, S. Bauer, N. Washington, M. Westerfield, G. Gkoutos, P. Schofield, D. Smedley, S. E. Lewis, P. N. Robinson, C. J. Mungall, Construction and accessibility of a cross-species phenotype ontology along with gene annotations for biomedical research. *F1000Research.* **2**, 30 (2013).

49. E. Turro, S.-Y. Su, Â. Gonçalves, L. J. M. Coin, S. Richardson, A. Lewin, Haplotype and isoform specific expression estimation using multi-mapping RNA-seq reads. *Genome Biol.* **12**, R13 (2011).

50. E. Turro, W. J. Astle, S. Tavaré, Flexible analysis of RNA-seq data using mixed effects models. *Bioinformatics.* **30**, 180–8 (2014).

51. M. Di Michele, C. Thys, E. Waelkens, L. Overbergh, W. D'Hertog, C. Mathieu, R. De Vos, K. Peerlinck, C. Van Geet, K. Freson, An integrated proteomics and genomics analysis to unravel a heterogeneous platelet secretion defect. *J. Proteomics.* **74**, 902–13 (2011).

52. C. Hermans, C. Wittevrongel, C. Thys, P. A. Smethurst, C. Van Geet, K. Freson, A compound heterozygous mutation in glycoprotein VI in a patient with a bleeding disorder. *J. Thromb. Haemost.* **7**, 1356–63 (2009).

53. A. Bonnefoy, K. Daenens, H. B. Feys, R. De Vos, P. Vandervoort, J. Vermylen, J. Lawler,

M. F. Hoylaerts, Thrombospondin-1 controls vascular platelet recruitment and thrombus adherence in mice by protecting (sub)endothelial VWF from cleavage by ADAMTS13. *Blood.* **107**, 955–64 (2006).

54. M. A. Seeliger, B. Nagar, F. Frank, X. Cao, M. N. Henderson, J. Kuriyan, c-Src binds to the cancer drug imatinib with an inactive Abl/c-Kit conformation and a distributed thermodynamic penalty. *Structure.* **15**, 299–311 (2007).

55. S. Louwette, L. Régal, C. Wittevrongel, C. Thys, G. Vandeweeghde, E. Decuyper, P. Leemans, R. De Vos, C. Van Geet, J. Jaeken, K. Freson, NPC1 defect results in abnormal platelet formation and function: studies in Niemann-Pick disease type C1 patients and zebrafish. *Hum. Mol. Genet.* **22**, 61–73 (2013).

56. K. Freson, K. Peeters, R. De Vos, C. Wittevrongel, C. Thys, M. F. Hoylaerts, J. Vermylen, C. Van Geet, PACAP and its receptor VPAC1 regulate megakaryocyte maturation: therapeutic implications. *Blood.* **111**, 1885–93 (2008).

57. A. Ibrahimi, G. Vande Velde, V. Reumers, J. Toelen, I. Thiry, C. Vandeputte, S. Vets, C. Deroose, G. Bormans, V. Baekelandt, Z. Debyser, R. Gijsbers, Highly efficient multicistronic lentiviral vectors with peptide 2A sequences. *Hum. Gene Ther.* **20**, 845–60 (2009).

58. S. Louwette, C. Van Geet, K. Freson, Regulators of G protein signaling: role in hematopoiesis, megakaryopoiesis and platelet function. *J. Thromb. Haemost.* **10**, 2215–22 (2012).

59. C. Rooryck, A. Diaz-Font, D. P. S. Osborn, E. Chabchoub, V. Hernandez- Hernandez, H. Shamseldin, J. Kenny, A. Waters, D. Jenkins, A. Al Kaissi, G. F. Leal, B. Dallapiccola, F. Carnevale, M. Bitner-Glindzicz, M. Lees, R. Hennekam, P. Stanier, A. J. Burns, H. Peeters, F. S. Alkuraya, P. L. Beales, Mutations in lectin complement pathway genes COLEC11 and MASP1 cause 3MC syndrome. *Nat. Genet.* **43**, 197–U36 (2011).

60. D. Bates, M. Mächler, B. M. Bolker, S. C. Walker, Fitting Linear Mixed-Effects Models using lme4. *J. Stat. Softw. …,* 1–51 (2014).

ACKNOWLEDGEMENTS

TABLES
Table 1. Clinical phenotype and blood and bone marrow analysis results of four affected cases (13, 19, 31 and 35) and unaffected case 21 with enhanced assessment.

FIGURES
Fig. 1. Selection of the c.1579G>A mutation in *SRC* as a candidate pathogenic variant. (**A**) Pedigree showing male (square) and female (circle) members who have macrothrombocytopenia (blue), are unaffected (empty) or are without clinical information (gray), some of whom are deceased (slash). Cases carry the c.1579G>A mutation in *SRC* (M) in variant calling done by whole genome sequencing (#), whole exome sequencing (✱) and/or Sanger sequencing (all genotyped subjects). (**B**) Barplot showing the mean phenotypic similarity of cases 13, 31 and 35 to OMIM/MPO phenotypes associated with each gene, truncated at the top 20 genes, with novel variants absent from control data indicated by * and *SRC* highlighted in blue (as in (**D**) and (**E**)). (**C**) The numbers inside each node indicate which cases were coded with the corresponding HPO term. Terms in green are also present in the OMIM/MPO entries for *SRC*/*Src*. The size of each node is determined by its contribution to the mean phenotype similarity score between the three cases and the OMIM/MPO terms for

SRC/Src. See materials and methods for abbreviations. The cases of this pedigree were the only ones enrolled in the BRIDGE Bleeding and Platelet Disorders study coded with "Thrombocytopenia", "Myelofibrosis" and "Abnormality of the skeletal system". (**D**) Barplot showing the CADD phred score of the rare variant for each candidate gene. (**E**) The probability that each candidate gene is specifically over-expressed in MKs compared to blood stem cells and six other hematopoietic progenitors.

Fig. 2. Platelet phenotype related to SRC E527K gain-of-function mutation. (**A**) Blood smears from cases 31, 35 and 19 showing large platelets with a grayish appearance (black arrows). EM images of platelets for the three cases show (i) round platelets that rarely have a normal discoid shape, (ii) some dysmorphic platelets (DP), (iii) platelets with many open canalicular system (OCS)-forming vacuoles, (iv) a reduced level of microtubules (MT), (v) a reduced number of α-granules with a subpopulation of platelets exhibiting small granules (SG) or a near absence of all types of granules, reduced levels of internal membranes and a cytoplasm that appears amorphous. Horizontal bars indicate 1μm. See fig. S2 for EM images of normal platelets. (**B,C**) The platelet area and number of total alpha granules/platelet were quantified. Values are the mean and s.e.m as quantified for 50 randomly selected platelets for a healthy control (C) and cases 31, 35 and 19. One-way ANOVA with Bonferroni's multiple test**.** ****$p<0.0001$ (**D**) Western blot shows Vwf, Tsp1 and $\beta3$ integrin (ITGB as loading control) expression for platelets from a healthy control (C) and cases 31, 35 and 19. (**E**) **Left panel**. Flow cytometric analysis of P-selectin (CD62P) on the outer-membrane of platelets from case 35 (blue) and an unrelated healthy control (red, WT) after activation with collagen. **Right panel**. Summarised mean fluorescence intensities (MFI) for two healthy controls (red) and cases 19, 31 and 35 (blue) (see Supplementary Materials for analysis and fig. S3 for representative image of convulxin activation). (**F**) Crystallography-based models of the SRC structures show the interaction of the C-terminal regulatory tail with SH2 domain residues (fig. S5). **Left panel.** Zoom-in of the surroundings of the atomic environment of phosphorylated (P)-Y530 in wild-type SRC. Residues 158, 159, 178 and 181 of the SH2 domain involved in the interaction with residues 520, 521, 527 and 530 of the kinase domain (Velcro-strap) are in stick mode. Electrostatic interactions are shown with orange dotted lines and the temperature factor of the Velcro-strap is shown with a blue to red gradient (from lower and higher flexibility, respectively). **Right panel.** Similar structure for the SRC-K527 mutant protein but note the reduced predicted flexibility of the Velcro-strap (residues are blue) leading to a rearrangement between the Velcro-strap and SH2 domain. (**G**) **Left panel**. SRC kinase activity was measured with a colorimetric ELISA assay using different concentrations of ATP. The optical density at 450 nm (OD_{450}) was measured for GST-tagged wild type (GST-WT) and E527K (GST-527K) SRC protein with and without the addition of the SRC inhibitor. **Right panel**. Total (Pan), active (Y419) and inactive (Y530) SRC protein levels were assayed via immunoblot analysis of GST-WT and GST-527K fusion proteins used for the kinase assay.

Fig. 3. *SRC* **expression studies in platelets and transfected COS-7 cells.** (**A**) Triplicate western blots for total (Pan), active (Y419) and inactive (Y530) SRC were performed for platelet lysates obtained from non-stimulated and collagen-stimulated platelets for 30, 60 and 300 seconds for 2 controls and 3 cases (19, 31 and 35). **Left panel**. All blots were quantified and we fitted a linear mixed model to the log normalized mean pixel intensity (MPI) data of the SRC protein levels in platelets for carriers of the wild-type or mutant allele. The residuals after adjusting for genotype and antibody fixed effects are shown for the Y419- and Y530-targeting antibodies. **Right panel.** Representative blot. GAPDH was used as normalisation control. The unaffected case 21 has normal levels of total, active and inactive SRC in non-stimulated platelets. (**B**) Immunoblot analysis showing whole-platelet tyrosine phosphorylation (4G10 antibody) stimulated by collagen or ADP from 30 to 300 seconds in platelets from a control and cases 35 and 19. The same platelet lysates were used as above. (**C**) Western blots were performed for lysates obtained from COS-7 cells transfected with empty (V), SRC wild-type (WT) and SRC-K527 vectors from a triplicate transfection experiment. **Left panel.** Residuals obtained after fitting a model without an interaction term between presence of the mutant allele and the level of binding of anti-Y530 antibody. Such a term is required to obtain a good model fit due to the dramatic decrease of Y530 in cells carrying the mutant rather than the WT allele (see supplementary materials). **Right panel.** Representative blot. (**D**) Western blots were performed for lysates obtained from COS-7 cells transfected with the same SRC vectors as in (C) with or without addition of SRC inhibitor-1 (inh) in a triplicate transfection experiment. **Left panel.** Residuals obtained after fitting a model without a fixed inhibitor term. Such a term is required to obtain a good model fit (see supplementary materials). **Right panel.** Representative blot.

Fig. 4. Effect of the SRC E527K gain-of-function mutation on megakaryopoiesis. (**A**) Plasma TPO levels (pg/ml) and platelet count (PLT) (x 10^9/L) for a patient with proven congenital amegakaryocytic thrombocytopenia (CAMT) (black square), three SRC cases (blue symbols) and five unrelated healthy controls (red circles). The assay sensitivity limit is 25 pg/ml (indicated by red line). Squares and circles in blue represent samples collected in 2004 and 2014, respectively for SRC cases . (**B**) Total amount of CFU-MK (upper panel) and CFU-GEMM (lower panel) colonies derived from peripheral blood CD34[+] mononuclear cells from a control (C1) and case 31 counted at days 7 and 12 of culture. Values are the mean and s.e.m as quantified for a triplicate plating experiment. One-way ANOVA with Bonferroni's multiple test**.** Upper panel Day 7 **p=0.0059, Day 12 ***p=0.0006 Lower panel Day 7 **p=0.0015, Day 12**p=0.0048. (**C**) **Left panel**. Representative light microscopy images of cultured MK showing formation of proplatelet extensions (white arrows) for the controls (C1 and C2). Proplatelet forming MKs are almost absent for the cases. Scale bars, 20 μm. **Right panel.** Proplatelet formation was quantified in MK suspension triplicated liquid cultures (for each condition) performed on two separate occasions (firstly, two controls and three cases and, secondly, a control, case 31 and unaffected case 21 with and without SRC inhibitor-1). The proportion of proplatelet formation was lower in the cultures from cases compared to controls but normal for the unaffected case. Addition of the inhibitor resulted in a rescue of the proplatelet formation

defect. Values are the mean and s.e.m as counted for MKs present in 10 randomly selected slides for each condition. One-way ANOVA with Bonferroni's correction. ****p<0.0001, *p=0.049. (**D**) Representative immunofluorescence confocal microscopy images of differentiated fibrinogen-adhered MKs at day 12 of culture visualised for active SRC (green, Y419) and DAPI (blue), showing rosette-like structures in MKs from case 31 while control MK (C3) only stained weakly positive for active SRC. Scale bars, 5 µm. Quantification of MKs positive for active SRC staining showed a significant difference between controls and cases. Values are the mean and s.e.m as counted for 10 randomly selected images for each condition. *t*-test. *p* = 0.0003 (**E**) Western blots for total (Pan), active (Y419), inactive (Y530) SRC, GFP and overall tyrosine phosphorylation (4G10) were performed for lentiviral-transduced cells differentiated for 12 days to MKs. GFP only generates a signal for the control vector. WT-transduced cells overexpress inactive SRC while E527K-transduced cells express active SRC. The overexpression of SRC in WT and E527K-transduced cells is also visible in the 4G10 blots though E527K-transduced MKs also display increased tyrosine phosphorylation of other proteins. (**F**) Quantification of the ploidy analysis of two independent transduction experiments showed a significant left shift for E527K transduced MKs (blue) indicative of more immature diploid (2N DNA) MKs compared to WT (red). Two-way ANOVA. For the interaction between ploidy and genotype ****p<0.0001 (**G**) Proplatelet formation was quantified in MK suspension cultures from a duplicated transduction experiment in a blinded fashion. The proportion of proplatelet formation was lower in the cultures from E527K transduced cells compared to WT conditions. Values are the mean and s.e.m as counted for MKs present in 10 randomly selected slides for each condition. One-way ANOVA with Bonferroni's correction. For both WT-1 vs E527K-1 and WT-2 vs E527K-2: ****p<0.0001. (**H**) Immunofluorescence confocal microscopy images of differentiated fibrinogen-adhered and lentiviral-transduced MKs at day 12 of culture visualised for active SRC (green, Y419) and phalloidin (red, F-actin), showing co-localisation in MKs from E527K-transduced MKs but not in WT-transduced cells. Scale bars, 20 µm.

Fig. 5. Phenotype analysis of zebrafish injected with buffer (C), *src*-E525K or *src*-WT mRNA in the absence or presence of the SRC inhibitor-1 (inh). (**A**) **Left panel.** Stereomicroscope images of the caudal hematopoietic tissue (CHT) region in the tail at 3 dpf to visualize the GFP-labeled thrombocytes using *Tg(cd41:EGFP)* zebrafish. **Middle panel.** Immunoblot analysis of 15 lysed embryos at 3 dpf for expression of total Src, GFP and total tyrosine phosphorylation using 4G10 antibody. **Right panel.** Blot quantification of a triplicate injection experiment showed a trend for reduced GFP levels for E525K injected embryos that is corrected when the inhibitor is added to the fish water. (B) Alcian blue cartilage staining at 5dpf showing cartilage defects in *src-E525K* overexpression embryos that can be rescued by SRC inhibition. pq, palatoquadrate; mc, Meckel's cartilage; cb, ceratobranchials; cl, cranial length as specified in fig. S15. **Left panel.** Representative images. **Middle panel.** Graph of length ratio of pq, mc and cb cartilages standardized by cranial length. **Right panel.** Graph of area ratio of pq, mc and cb cartilages standardized by cranial length. Values are the mean and s.e.m as quantified for 6 randomly selected embryos for each condition. One-way ANOVA with

Bonferroni's correction. Length: ***p=0.0002, ****p<0.0001 and *p=0.0196. Area: ****p<0.0001, ***p=0.0006