REPORT

A Recurrent *PDGFRB* Mutation Causes Familial Infantile Myofibromatosis

Yee Him Cheung,[1,20] Tenzin Gayden,[2,20] Philippe M. Campeau,[3] Charles A. LeDuc,[1] Donna Russo,[1] Van-Hung Nguyen,[4] Jiancheng Guo,[1] Ming Qi,[5,6,7] Yanfang Guan,[7] Steffen Albrecht,[4] Brenda Moroz,[8] Karen W. Eldin,[9] James T. Lu,[10,11] Jeremy Schwartzentruber,[12] David Malkin,[13] Albert M. Berghuis,[14] Sherif Emil,[15] Richard A. Gibbs,[10] David L. Burk,[14] Megan Vanstone,[16] Brendan H. Lee,[3,17] David Orchard,[18] Kym M. Boycott,[16] Wendy K. Chung,[1,*] and Nada Jabado[2,19,*]

Infantile myofibromatosis (IM) is the most common benign fibrous tumor of soft tissues affecting young children. By using whole-exome sequencing, RNA sequencing, and targeted sequencing, we investigated germline and tumor DNA in individuals from four distinct families with the familial form of IM and in five simplex IM cases with no previous family history of this disease. We identified a germline mutation c.1681C>T (p.Arg561Cys) in platelet-derived growth factor receptor β (*PDGFRB*) in all 11 affected individuals with familial IM, although none of the five individuals with nonfamilial IM had mutations in this gene. We further identified a second heterozygous mutation in *PDGFRB* in two myofibromas from one of the affected familial cases, indicative of a potential second hit in this gene in the tumor. PDGFR-β promotes growth of mesenchymal cells, including blood vessels and smooth muscles, which are affected in IM. Our findings indicate p.Arg561Cys substitution in PDGFR-β as a cause of the dominant form of this disease. They provide a rationale for further investigations of this specific mutation and gene to assess the benefits of targeted therapies against PDGFR-β in aggressive life-threatening familial forms of the disease.

Infantile myofibromatosis (IM) (MIM 228550) is the most common benign tumor of soft tissue of infancy and childhood.[1] First described by Stout,[2] IM is characterized by solitary or multiple nodules in the skin, muscle, subcutaneous tissues, bone, and occasionally viscera. IM is simplex or occurs with an autosomal-dominant (AD) mode of inheritance.[3,4] Myofibromas are usually present at birth or develop shortly thereafter, with 90% of cases occurring before the age of 2 years.[5] Solitary and multicentric IMs that do not involve the viscera tend to spontaneously regress and their recurrence is relatively low. However, multicentric IM with visceral involvement has a poor outcome, with a mortality rate greater than 70% despite aggressive therapies.[6,7] The molecular etiology of the disease remains unknown.

To determine the genetic defect(s) underlying IM and whether the causes of familial and simplex IM are similar, we studied 11 individuals from 4 IM-affected families and 5 simplex cases. The clinical features and genotypes of the individuals investigated in this study are presented in Table S1 (available online) and the pedigrees of the four families are shown in Figure 1. The studies were approved by the Institutional Review Boards of Columbia University, the Baylor College of Medicine, McGill University Health Centre Research Institute, and the Children's Hospital of Eastern Ontario. Blood and tumor samples were obtained with informed consent from the patients and their parents according to Canadian and US laws. Genomic DNA was isolated from blood and from frozen and paraffin-embedded tissues. Total RNA was extracted from tumor tissue excised from the abdominal wall of individual III-1 of family 2 (Figure 1).

We first focused on familial cases and performed next-generation sequencing on DNA and RNA extracted from a discovery set of IM-affected familial cases. Whole-exome sequencing (WES) was performed on germline DNA from two affected siblings from a family of Chinese origin (family 1, Figure 1). The brother carried the typical solitary form, and the sister was treated for a visceral type with multiple myofibromas of the orbit and supranasal region. Exomes were captured with the Illumina TruSeq kit and were sequenced on an Illumina Hiseq 2000 with 100 bp

[1]Department of Pediatrics, Columbia University, New York, NY 10032, USA; [2]Department of Human Genetics, McGill University, Montreal, QC H3A 1B1, Canada; [3]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA; [4]Department of Pathology, Montreal Children's Hospital, McGill University Health Center, Montreal, QC H3H 1P3, Canada; [5]Center for Genetic and Genomic Medicine, Zhejiang University Medical School, Zheijang 310058, China; [6]Department of Pathology and Laboratory Medicine, University of Rochester, Rochester, NY 14642, USA; [7]BGI-Shenzhen, Shenzhen 518083, China; [8]Department of Pediatric Dermatology, Montreal Children's Hospital, Montreal, QC H3H 1P3, Canada; [9]Department of Pathology and Immunology, Baylor College of Medicine, Houston, TX 77030, USA; [10]Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX 77030, USA; [11]Department of Structural and Computational Biology & Molecular Biophysics, Baylor College of Medicine, Houston, TX 77030, USA; [12]McGill University and Genome Quebec Innovation Centre, Montreal, QC H3A 1A4, Canada; [13]Division of Hematology/Oncology, The Hospital for Sick Children, Department of Pediatrics, University of Toronto, Toronto, ON M5G 1X8, Canada; [14]Department of Biochemistry and Groupe de Recherche Axé sur la Structure des Protéines, McGill University, Montreal, QC H3G 0B1, Canada; [15]Department of Pediatric Surgery, McGill University, Montreal Children's Hospital, Montreal, QC H3H 1P3, Canada; [16]Children's Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, ON K1H 8L1, Canada; [17]Howard Hughes Medical Institute, Houston, TX 77030, USA; [18]Department of Dermatology, Royal Children's Hospital, Melbourne, Victoria 3052, Australia; [19]Department of Pediatrics, McGill University and the McGill University Health Center Research Institute, Montreal, QC H3H 2Z3, Canada
[20]These authors contributed equally to this work
*Correspondence: wkc15@mail.cumc.columbia.edu (W.K.C.), nada.jabado@mcgill.ca (N.J.)
http://dx.doi.org/10.1016/j.ajhg.2013.04.026. ©2013 by The American Society of Human Genetics. All rights reserved.

Family 1

Family 2

Family 3

Family 4

In both families, we identified the same heterozygous c.1681C>T (p.Arg561Cys) variant in *PDGFRB* (MIM 173410; RefSeq accession number NM_002609.3), which encodes a receptor tyrosine kinase (RTK). The missense variant was predicted to be damaging by all prediction algorithms and affects a highly conserved amino acid residue. This mutation was absent from our control in-house exome database, dbSNP135, the 1000 Genomes project,[10] and NHLBI Exome Sequencing Project. Sanger sequencing (Table S2) confirmed the heterozygous mutation in all affected members of both families 1 and 2. Because the mutation was absent in family 2 members I-1 and I-2, we determined that it was acquired de novo in II-2. In the same individual, we identified a second variant, c.1998C>A (p.Asn666Lys), that is predicted to be damaging. This variant was identified in a second gingival myofibroma but was absent from a neck myofibroma and from blood, indicating that the variant is somatic. Neither of these two genetic alterations was present in nonaffected family members (Figure 1) or in the 43 somatic missense mutations of *PDGFRB* listed in the Catalogue Of Somatic Mutations In Cancer.[15]

We then screened for mutations in *PDGFRB* by using Sanger sequencing in an additional two separate families affected with IM (Figure 1, families 3 and 4) and in five simplex IM cases, including three with visceral IM (Table S1). Family 3 is of French-Canadian origin and the two affected siblings (III-2 and III-3) had multiple myofibromas of the skin on the face and upper arms that spontaneously regressed at the age of 4 years, similar to the father who also had myofibroma that spontaneously resolved. Family 4 is of European ancestry and was previously reported.[16] The two affected brothers (II-1 and II-3) were diagnosed with multiple myofibromas at the young ages of 3 and 11 weeks, respectively, with no clinical evidence of visceral involvement. The mother had swelling in the left side of her neck at 7 months of age, which was found to be fibromatosis and subsequently excised. The five simplex cases included in this study were taken from resected myofibromas from children at the following ages and locations: at birth (cheek), 34 days (pelvic myofibroma, reported as case 2 in a previous study[17]), 13 days (lesion from right hemidiaphragm), 5 days (visceral myofibroma),

paired-end reads. Reads were aligned against the reference human genome (UCSC Genome Browser hg19) with BWA,[8] variants called and annotated as previously described.[9] Given the rarity of the disease, we eliminated variants with minor allele frequency (MAF) greater than 1% in the 1000 Genomes[10] and NHLBI GO Exome Sequencing Project databases or greater than 5% in approximately 500 exomes previously sequenced at our center. We also performed RNA-seq on an abdominal wall myofibroma from the child (III-1) of an affected mother-child pair of European ancestry in family 2 (Figure 1). Both of these affected individuals suffer from multiple myofibromas of the head, neck, and abdominal wall, which were either surgically resected or spontaneously regressed. In brief, mRNAs were enriched from total RNA with poly(A) selection followed by library preparation by Illumina TruSeq RNA prep kit and sequencing on Illumina HiSeq 2000 with single-end 100 bp reads. The pass filter reads were then mapped to the reference human genome (NCBI build 37) by TopHat[11] (v.1.3.3). For each read, up to two mismatches and ten multiple hits were allowed during the mapping. Variants were called with SAMtools (v.0.1.17), mpileup, and bcftools, filtered by mapping quality ≥ 5, read depth ≥ 5, and base quality ≥ 17. Functional annotations were obtained by SeattleSeq Annotation 134 (NCBI and CCDS 2011) and ANNOVAR.[12] The RNA-seq data revealed a total of 28,141 SNVs and 923 short indels in 6,838 genes. Assuming AD inheritance, variants on the X and Y chromosomes and mitochondrial genome were excluded. We retained only nonsynonymous and splice variants and eliminated variants with MAF greater than 1% in the 1000 Genomes database.[10] We also removed variants predicted to be tolerated, benign, or unknown by both SIFT[13] and PolyPhen-2.[14] After the filtering procedures, there were 385 SNVs and 43 indels in the remaining 338 genes.

and 1 month (femur and lung). Matched blood was available in two cases and was also analyzed for *PDFGRB* mutations. All of the affected individuals with familial IM harbored the same missense variant, c.1681C>T, thereby confirming the role of *PDGFRB* in the pathogenesis of familial IM as well as the de novo occurrence of this mutation in a family member (II-2 in family 2 and II-1 in family 3, Figure 1). None of the five individuals with simplex IM harbored either of the two *PDGFRB* mutations in myofibromas or in germline DNA in the two cases for which this material was also available.

PDGFRB encodes the β polypeptide of the platelet-derived growth factor receptor (PDGFR-β), a RTK and a mitogen for mesenchyme-derived cells, including fibroblasts and smooth muscle,[18,19] which are targeted in IM. Mouse knockout studies showed the critical role of PDGFR-β signaling in the embryonic development of specific subsets of smooth muscle cells,[20,21] particularly in the proper recruitment of vascular smooth muscle cells and pericytes to developing blood vessel walls.[16–19] Deregulation of PDGF signaling (usually through somatic mutations, overexpression, or fusion with other genes) has been associated with several human disorders and cancers affecting cells of mesenchymal origin (reviewed in Andrae et al.,[18] Trojanowska,[22] and George[23]), and more recently, p.Leu658Pro and p.Arg987Trp *PDGFRB* germline variants have been associated with idiopathic basal ganglia calcification (MIM 615007).[24]

The germline mutation in *PDFGRB* we identify herein has not been previously reported. To gain insight into its potential effects on receptor activity, we constructed a homology model of the structure of the cytoplasmic domain of human PDGFR-β. The sequence of the PDGFR-β cytoplasmic domain (residues 554–1106) was obtained from the UniProt database and converted to PIR format for use in the Modeler program suite.[25] A search for related sequences of known structure identified the autoinhibited form of human KIT kinase (PDB ID 1T45) as a useful template upon which to model the structure of PDGFR-β. KIT and PDGFR-β are members of the same receptor tyrosine kinase family (class III) and possess a high degree of sequence identity in the cytoplasmic region (~57%). Because only the extracellular domain of PDGFR-β has been previously modeled, the human KIT kinase was used for this modeling, based on extensive homology with PDGFRs.[26] The sequence of the PDGFR-β cytoplasmic domain was aligned to that of KIT and several three-dimensional models of PDGFR-β were generated with Modeler according to this target-template alignment. The resulting structural models were reasonable, as suggested by the scores from Modeler and by visual inspection with the program PyMOL.[27] The model includes the cytoplasmic juxtamembrane (JM) domain and the two lobes of the kinase domain. The structures of all models were highly similar, except in the area of the kinase insert segment (approximately residues 700–795). Because the KIT template structure lacks the kinase insert segment found in PDGFR-β,

modeling of these residues is less reliable. However, the c.1681C>T (p.Arg561Cys) and c.1998C>A (p.Asn666Lys) alterations are situated far from the site of the kinase insert and consequently are not impacted by uncertainty for that segment of the PDGFR-β model. To assess the role of residues Arg561 and Asn666 and the potential impact of the observed mutations for the function of the PDGFR-β JM and kinase domains, we compared the PDGFR-β model with the structures of the autoinhibited (PDB ID 1T45) and active (PDB ID 1PKG) forms of KIT.

Arg561 maps to the JM region of PDGFR-β between the helical transmembrane segment and the kinase domain, lying at the boundary of the JM domain and the short section of polypeptide that links it to the transmembrane helix. The JM domain is autoinhibitory, masking the catalytic cleft when the receptor is not bound by its ligand.[28] In the model, the side chain of Arg561 forms a salt bridge with residue Glu644 in an adjacent α helix from the C-terminal lobe of the kinase domain (Figure 2C). This interaction would be expected to tether this part of the JM domain to the C-terminal lobe and contribute to the binding of the autoinhibitory JM segment to the kinase domain. A p.Arg561Cys change would abrogate the salt bridge, weakening the binding of the JM domain. This would compromise the autoinhibitory role of the JM domain in preventing receptor firing at baseline. The second somatic change, c.1998C>A (p.Asn666Lys), is located in the N-terminal lobe of the kinase domain, near the binding site for RTK inhibitors such as imatinib and sunitinib. In the PDGFR-β model, the side chain of Asn666 participates in hydrogen-bonding interactions with the backbone of His661. A similar interaction is observed between Asn655 and His650 in the structure of the autoinhibited form of KIT kinase.[29] Interestingly, the structure of the active form of KIT kinase is different in this area.[30] The side chain of Asn655 is oriented in a different direction and no longer interacts with residue H650 (Figure 2D). The p.Asn666Lys substitution in PDGFR-β would thus abolish the interaction between Asn666 and His661, altering the interactions in this area of the protein and possibly leading to a structure more similar to the active conformation of KIT kinase.

PDGFR-β signaling is critical for normal development of mesenchymal tissues.[21,22] IM affects infants during their first years of life when expression of PDGFR-β and its ligands are highest. Our findings indicate that a recurrent germline mutation, c.1681C>T (p.Arg561Cys), in PDGFRB is responsible for the autosomal-dominant familial but not for the simplex form of IM. The *PDGFRB* mutation in familial IM is predicted to decrease autoinhibition of the JM domain at baseline, leading to increased kinase firing and promoting the formation of myofibromas in tissues with high PDGFR-β signaling activity. This may explain the perinatal and early childhood prevalence of this disease and its regression with age. Aberrant signaling through PDGFR-β in the genesis of familial IM is further supported by the presence of a second hit in this RTK in visceral myofibroma. Several drugs effectively inhibit the
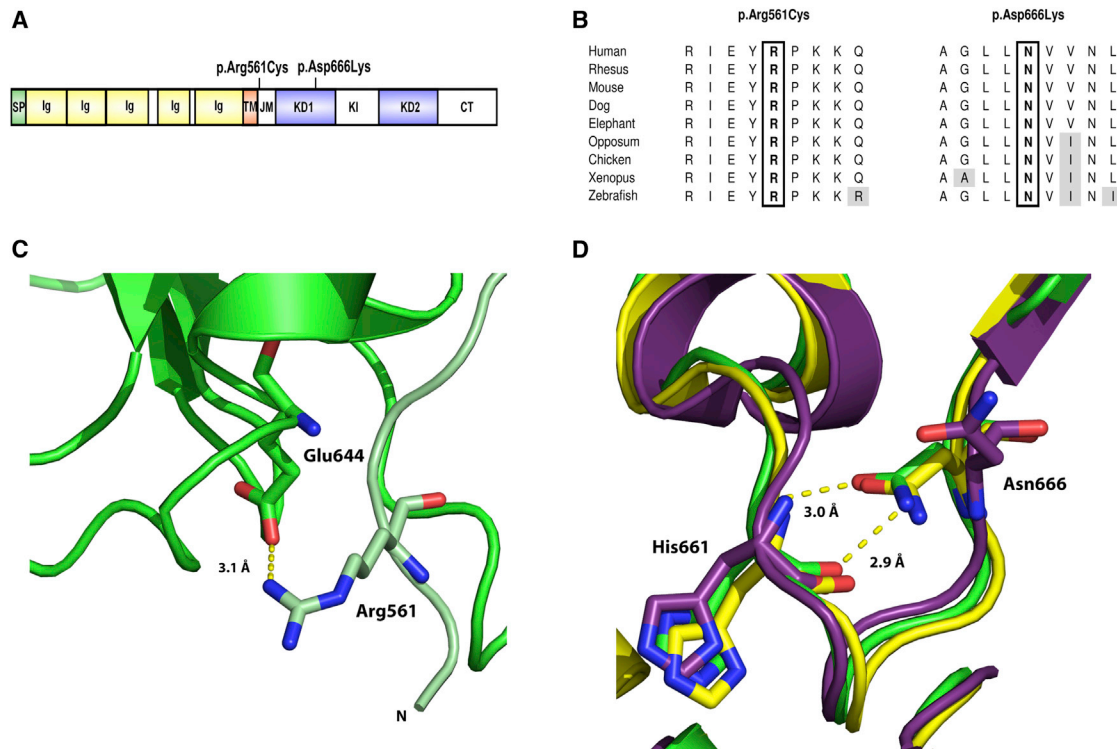
Figure 2. PDGFR-β Substitutions in Familial Infantile Myofibromatosis

(A) Domain structure of PDGFR-β and position of identified mutations. Abbreviations are as follows: SP, signal peptide; Ig, immunoglobulin-like C2 domain; TM, transmembrane domain; JM, juxtamembrane domain; KD1, first split kinase domain; KI, kinase insert; KD2, second split kinase domain; CT, C-terminal tail domain.

(B) Amino acid conservation across species for the mutated residues.

(C) A ribbon diagram showing the interaction between Arg561 and Glu644 in the PDGFR-β model. The backbone of the model is shown in green, with residues of the JM domain colored light green. The p.Arg561Cys substitution would abrogate this interaction.

(D) A ribbon diagram showing the interaction between Asn666 and His661 in the PDGFR-β model (green). The structure of the auto-inhibited (yellow) and active (purple) forms of KIT kinase are shown for comparison. A p.Asn666Lys change would abolish the interaction linking Asn666 and His661.

kinase activity of PDGFR-β, and our data support further investigations into these *PDGFRB* mutations to assess the potential use of these inhibitors in the treatment of life-threatening forms of the IM, where limited effective therapeutic options are currently available.

Supplemental Data

Supplemental Data include two tables and can be found with this article online at http://www.cell.com/AJHG/.

Web Resources

The URLs for data presented herein are as follows:

1000 Genomes, http://browser.1000genomes.org
ANNOVAR, http://www.openbioinformatics.org/annovar/
NCBI Gene, http://www.ncbi.nlm.nih.gov/gene
NHLBI Exome Sequencing Project (ESP) Exome Variant Server, http://evs.gs.washington.edu/EVS/

Online Mendelian Inheritance in Man (OMIM), http://www.omim.org/

PolyPhen-2, http://www.genetics.bwh.harvard.edu/pph2/

SeattleSeq Annotation 137, http://snp.gs.washington.edu/SeattleSeqAnnotation137/

SIFT, http://sift.bii.a-star.edu.sg/

UCSC Genome Browser, http://genome.ucsc.edu

UniProt, http://www.uniprot.org/

References

1. Wiswell, T.E., Davis, J., Cunningham, B.E., Solenberger, R., and Thomas, P.J. (1988). Infantile myofibromatosis: the most common fibrous tumor of infancy. J. Pediatr. Surg. *23*, 315–318.

2. Stout, A.P. (1954). Juvenile fibromatoses. Cancer *7*, 953–978.

3. Jennings, T.A., Duray, P.H., Collins, F.S., Sabetta, J., and Enzinger, F.M. (1984). Infantile myofibromatosis. Evidence for an autosomal-dominant disorder. Am. J. Surg. Pathol. *8*, 529–538.

4. Zand, D.J., Huff, D., Everman, D., Russell, K., Saitta, S., McDonald-McGinn, D., and Zackai, E.H. (2004). Autosomal dominant inheritance of infantile myofibromatosis. Am. J. Med. Genet. A. *126A*, 261–266.

5. Chung, E.B., and Enzinger, F.M. (1981). Infantile myofibromatosis. Cancer *48*, 1807–1818.

6. Auriti, C., Kieran, M.W., Deb, G., Devito, R., Pasquini, L., and Danhaive, O. (2008). Remission of infantile generalized myofibromatosis after interferon alpha therapy. J. Pediatr. Hematol. Oncol. *30*, 179–181.

7. Azzam, R., Abboud, M., Muwakkit, S., Khoury, N., and Saab, R. (2009). First-line therapy of generalized infantile myofibromatosis with low-dose vinblastine and methotrexate. Pediatr. Blood Cancer *52*, 308.

8. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

9. Schwartzentruber, J., Korshunov, A., Liu, X.-Y., Jones, D.T.W., Pfaff, E., Jacob, K., Sturm, D., Fontebasso, A.M., Quang, D.-A.K., Tönjes, M., et al. (2012). Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature *482*, 226–231.

10. Abecasis, G.R., Altshuler, D., Auton, A., Brooks, L.D., Durbin, R.M., Gibbs, R.A., Hurles, M.E., and McVean, G.A.; 1000 Genomes Project Consortium. (2010). A map of human genome variation from population-scale sequencing. Nature *467*, 1061–1073.

11. Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105–1111.

12. Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. *38*, e164.

13. Sim, N.-L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., and Ng, P.C. (2012). SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Res. *40*(Web Server issue), W452–W457.

14. Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. Nat. Methods *7*, 248–249.

15. Forbes, S.A., Bindal, N., Bamford, S., Cole, C., Kok, C.Y., Beare, D., Jia, M., Shepherd, R., Leung, K., Menzies, A., et al. (2011). COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. Nucleic Acids Res. *39*(Database issue), D945–D950.

16. Smith, A., and Orchard, D. (2011). Infantile myofibromatosis: two families supporting autosomal dominant inheritance. Australas. J. Dermatol. *52*, 214–217.

17. Wilson, M., Emil, S., Cowan, K., Kalechstein, S., Puligandla, P., Nguyen, V.-H., Laberge, J.-M., and Chou, S. (2013). Infantile myofibromas obstructing opposite ends of the gastrointestinal tract. J. Pediatr. Surg. *48*, 449–453.

18. Andrae, J., Gallini, R., and Betsholtz, C. (2008). Role of platelet-derived growth factors in physiology and medicine. Genes Dev. *22*, 1276–1312.

19. Fredriksson, L., Li, H., and Eriksson, U. (2004). The PDGF family: four gene products form five dimeric isoforms. Cytokine Growth Factor Rev. *15*, 197–204.

20. Levéen, P., Pekny, M., Gebre-Medhin, S., Swolin, B., Larsson, E., and Betsholtz, C. (1994). Mice deficient for PDGF B show renal, cardiovascular, and hematological abnormalities. Genes Dev. *8*, 1875–1887.

21. Soriano, P. (1994). Abnormal kidney development and hematological disorders in PDGF beta-receptor mutant mice. Genes Dev. *8*, 1888–1896.

22. Trojanowska, M. (2008). Role of PDGF in fibrotic diseases and systemic sclerosis. Rheumatology (Oxford) *47*(Suppl 5), v2–v4.

23. George, D. (2003). Targeting PDGF receptors in cancer—rationales and proof of concept clinical trials. Adv. Exp. Med. Biol. *532*, 141–151.

24. Nicolas, G., Pottier, C., Maltête, D., Coutant, S., Rovelet-Lecrux, A., Legallic, S., Rousseau, S., Vaschalde, Y., Guyant-Maréchal, L., Augustin, J., et al. (2013). Mutation of the PDGFRB gene as a cause of idiopathic basal ganglia calcification. Neurology *80*, 181–187.

25. Sali, A., and Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. *234*, 779–815.

26. Qiu, F.H., Ray, P., Brown, K., Barker, P.E., Jhanwar, S., Ruddle, F.H., and Besmer, P. (1988). Primary structure of c-kit: relationship with the CSF-1/PDGF receptor kinase family—oncogenic activation of v-kit involves deletion of extracellular domain and C terminus. EMBO J. *7*, 1003–1011.

27. Schrödinger, L. (2010). The {PyMOL} Molecular Graphics System, Version~1.3r1.

28. Hubbard, S.R. (2004). Juxtamembrane autoinhibition in receptor tyrosine kinases. Nat. Rev. Mol. Cell Biol. *5*, 464–471.

29. Mol, C.D., Dougan, D.R., Schneider, T.R., Skene, R.J., Kraus, M.L., Scheibe, D.N., Snell, G.P., Zou, H., Sang, B.-C., and Wilson, K.P. (2004). Structural basis for the autoinhibition and STI-571 inhibition of c-Kit tyrosine kinase. J. Biol. Chem. *279*, 31655–31663.

30. Mol, C.D., Lim, K.B., Sridhar, V., Zou, H., Chien, E.Y.T., Sang, B.-C., Nowakowski, J., Kassel, D.B., Cronin, C.N., and McRee, D.E. (2003). Structure of a c-kit product complex reveals the basis for kinase transactivation. J. Biol. Chem. *278*, 31461–31464.