Genetic sequence analysis of inherited bleeding diseases

Flora Peyvandi,1 Tom Kunicki,2 and David Lillicrap3

1Angelo Bianchi Bonomi Haemophilia and Thrombosis Centre, Fondazione IRCCS Ca’ Granda Ospedale Maggiore Policlinico, and Department of Pathophysiology and Transplantation, University of Milan, Luigi Villa Foundation, Milan, Italy; 2CHOC Children’s Hospital, Hematology Research, Orange, CA; and 3Department of Pathology and Molecular Medicine, Queen’s University, Kingston, Canada

The genes encoding the coagulation factor proteins were among the first human genes to be characterized over 25 years ago. Since then, significant progress has been made in the translational application of this information for the 2 commonest severe inherited bleeding disorders, hemophilia A and B. For these X-linked disorders, genetic characterization of the disease-causing mutations is now incorporated into the standard of care and genetic information is used for risk stratification of treatment complications. With electronic databases detailing >2100 unique mutations for hemophilia A and >1100 mutations for hemophilia B, these diseases are among the most extensively characterized inherited diseases in humans. Experience with the genetics of the rare bleeding disorders is, as expected, less well advanced. However, here again, electronic mutation databases have been developed and provide excellent guidance for the application of genetic analysis as a confirmatory approach to diagnosis. Most recently, progress has also been made in identifying the mutant loci in a variety of inherited platelet disorders, and these findings are beginning to be applied to the genetic diagnosis of these conditions. Investigation of patients with bleeding phenotypes without a diagnosis, using genome-wide strategies, may identify novel genes not previously recognized as playing a role in hemostasis. (Blood. 2013; 122(20):3423-3431)

The coagulation factor genes were among the first to be cloned and sequenced, beginning in the early 1980s. Now, in 2013, the detailed genetic structure of all of the identified constituents of the hemostatic process is well documented.

The purpose of this minireview is to highlight current knowledge and the potential translational utility of genetic sequence information for 3 categories of bleeding disease: (1) the common inherited bleeding disorders, hemophilia A, B, and von Willebrand disease (VWD); (2) the rare inherited coagulation factor deficiencies; and (3) inherited platelet disorders resulting in a bleeding phenotype. In light of the varying lengths of time since the initial genetic characterization of these various components of the hemostatic system, the extent of detailed information presented in this review is similarly variable.

Hemophilia A and B

The coagulation factor genes to be sequenced were the genes implicated in the common severe inherited bleeding conditions, hemophilia A and B.1,2 Very quickly, after the cloning of the F8 and F9 genes, strategies were initiated to incorporate this new information into hemophilia carrier testing and prenatal diagnosis.3

The 2 web-based mutation databases for the hemophilies now represent some of the most comprehensive documentation of the genetic landscape of inherited human disease. More than 2100 different F8 mutations have been associated with hemophilia A (http://hadh.org.uk/WebPages/PublicFiles/MutationSummary.htm), and more than 1100 F9 mutations have been identified in hemophilia B (http://www.factorix.org/) (See Table 1).

The sequences of both the F8 and F9 genes are minimally polymorphic, and the very few single-nucleotide variants present in the F8 gene have a variable frequency in different ethnic groups. This phenomenon has been proposed as a potential explanation for the differential propensity for FVIII inhibitor development between white and nonwhite populations, although this hypothesis awaits further confirmation.4

In hemophilia A, the association of a single recurrent gene inversion mutation in 45% of patients with a severe phenotype has facilitated genetic diagnostic approaches,5 but the incorporation of F8 sequencing to provide a more comprehensive evaluation of mutations is now also becoming a routine component of hemophilia management.6 In hemophilia B, no such common recurrent mutation exists, and a larger percentage of cases are the result of missense mutations and the subsequent production of a dysfunctional protein.7

In hemophilia A and B, genetic analysis of the F8 and F9 genes results in the identification of causative mutations in >95% of patients. This fact, along with the high likelihood of additional mutations being identified within introns8 and transcriptional regulatory regions9 (which are not included in the routine sequence analysis) suggest that locus heterogeneity for these traits is unlikely.

In the clinic, the identification of hemophilic mutations is currently used for 2 purposes: the definitive diagnosis of the carrier state and prenatal determination of the disease, and for incorporation into estimates for the risk of developing inhibitory antibodies to clotting factor replacement.10 The use of mutation detection for family counseling has been a dramatic improvement over previous coagulation test-based and linked polymorphic marker analysis for this purpose, and is now increasingly incorporated into routine clinical care. With regards to the genetic prediction of inhibitor development,
Although it was recognized that this treatment complication has a multifactorial etiology, the hemophilic genotype is currently the most influential contributory factor documented. Patients with multidomain deletions of either the F8 or F9 genes are at high risk of inhibitor development (50% to 80%), and in hemophilia B this risk is also associated with the potential for anaphylactic reactions to FIX concentrate infusion.11

In addition to its utility for clinical management, molecular genetic analysis of the hemophlias has also provided additional knowledge relating to structure/function and genotype/phenotype correlations (Figure 1). Two biologically interesting examples of these advances are summarized below.

### Hemophilia B Leiden

It has been recognized for several decades that there is a small group of hemophilia B patients who present with severe disease in childhood but whose factor IX levels increase progressively into the normal range following puberty. This hemophilia B Leiden phenotype has now been extensively characterized at the genetic level and found to be caused by a group of single-nucleotide substitutions clustered around the transcription start site of the factor IX gene.12 The postpubertal recovery of factor IX expression appears to be due principally to the activation of an androgen response element (ARE) approximately 30 nucleotides upstream of the transcription start site.13 Disruption of the ARE abolishes this transcriptional recovery.14 Although this unique recovery mechanism can be conveyed as positive information to families in which Leiden mutations are present, there are also reports suggesting that activation of the ARE is not the only explanation for this recovery mechanism.15,16

## Factor VIII assay–discrepant genotypes

During the past few years, there has been a growing realization that some patients with mild and moderate hemophilia A demonstrate significant discrepancies in the 2 laboratory assays used to quantify FVIII procoagulant function: the more frequently used 1-stage FVIII assay and the chromogenic factor VIII assay.17 In some instances, the magnitude of this discrepancy is sufficient to mask the diagnosis of hemophilia A. Genetic analysis of these cases has now revealed that the cause of this phenotype is missense substitutions localized to the interfaces between the factor VIII A domains that result in increased instability of the molecule.18 It is this instability phenotype that manifests through the assay variability described above relating principally to the different incubation times being employed in these laboratory tests (usually longer incubation times and thus better detection of instability mutants with chromogenic FVIII assays).

### VWD

The von Willebrand factor (VWF) gene was cloned just a year after FVIII in 1985.19-22 In contrast to the F8 locus, the VWF sequence is highly polymorphic (Table 1), and sequence analysis is further complicated by the presence of a partial VWF pseudogene on chromosome 22 that replicates the VWF sequence between exons 23 and 34 with 3% variance.23

Although diagnostic genetic sequence analysis for the hemophlias was implemented rapidly following the cloning of the F8 and F9 genes, the diagnostic application of sequencing of the VWF gene has been more complicated. This is due in part to the challenges in performing robust phenotypic testing for VWD, with the identification and differentiation of some VWD subtypes being especially problematic (eg, type 1 and some type 2 VWD variants). These phenotypic challenges led to an initial hope that genetic sequencing of VWF might provide a more reproducible and definitive

![Genetic Testing for Inherited Bleeding Disorders](image)

**Figure 1. Examples of the role of molecular diagnostic testing in the clinical management of inherited bleeding disorders.** In most instances, the initial diagnosis of these traits will still be made through the application of standard coagulation-based functional assays. The increasing utility of molecular testing is to provide additional information to aid in the future clinical management of these conditions.

<p>| <strong>Table 1. Genetic characteristics of the 3 common inherited bleeding disorders: hemophilia A, hemophilia B, and VWD</strong> |</p>
<table>
<thead>
<tr>
<th><strong>Disease</strong></th>
<th><strong>Gene</strong></th>
<th><strong>Mutation database</strong></th>
<th><strong>Mutation detection rate</strong></th>
<th><strong>VWF</strong> mutation detection rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemophilia A</td>
<td>F8</td>
<td><a href="http://www.hadb.org.uk">www.hadb.org.uk</a></td>
<td>2107 unique F8 mutations</td>
<td>~95%</td>
</tr>
<tr>
<td>Hemophilia B</td>
<td>F9</td>
<td><a href="http://www.factorix.org">www.factorix.org</a></td>
<td>1,106 unique F9 mutations</td>
<td>~95%</td>
</tr>
<tr>
<td>VWD</td>
<td>VWF</td>
<td><a href="http://www.vwf.group.shef.ac.uk">www.vwf.group.shef.ac.uk</a></td>
<td>399 unique VWF mutations</td>
<td>~65%</td>
</tr>
</tbody>
</table>

Databases accessed May 2013.
characterization of the various VWD subtypes. Investigations over the past decade have shown that the promise of “genetic clarification” has been only partially fulfilled.

There are clear examples where VWF genetic studies have been very successful in providing new or supplementary information to that derived from phenotypic testing. The prenatal diagnosis of type 3 VWD is the most obvious advance made possible by VWF sequencing. All causative mutations for this severe phenotype are located at the VWF locus, and their assignment as pathological variants is usually uncomplicated. With this information in mind, definitive genetic ascertainment of this phenotype can be offered for either chorionic villus or amniocyte samples early in pregnancy.

The other diagnostic area where sequence analysis has proven to be beneficial concerns the identification of type 2 variant forms of VWD. The confident phenotypic determination of type 2 variants is often difficult, and the localization of missense mutations to distinct regions of the VWF gene has especially facilitated identification of type 2B, 2M, and 2N subtypes. In each of these instances, confirmation of the diagnosis is important for optimal therapeutic management, and for both type 2B and 2N disease, where genotypes exist (platelet-type VWD and mild/moderate hemophilia A, respectively), this is particularly important. Where genetic sequencing has resulted in the generation of more questions than answers is in the investigation of type 1 VWD, the common mild/moderate reduction of functionally normal VWF. Ironically, for this VWD subtype in which phenotypic confirmation can be especially frustrating, the results of genetic analysis have, in general, only added to the complexity of the diagnostic challenge.

Four multicenter genetic studies of type 1 VWD have now been completed involving the sequence analysis of more than 500 index cases. The consistent finding in these studies has been that candidate VWF mutations have been identified in only approximately 65% of index cases. Even more perplexing is the fact that some of the candidate mutations (the majority of which are missense substitutions) have now been identified in individuals from different ethnic groups with no history of bleeding problems. Some of these variants have been studied in vitro using heterologous cell models, and very recently, VWF biosynthetic, storage, and secretion defects have been demonstrated with endothelial progenitor cells derived from type 1 VWD patients. Furthermore, in some type 1 patients, accelerated VWF clearance is the predominant pathogenic mechanism, and this phenomenon may also contribute more widely to this quantitative trait.

The question of the location of the other 35% of type 1 VWD mutations currently remains unanswered. The VWF introns and distal transcriptional regulatory sequences have yet to be analyzed in this population, so there is a possibility that at least some of these missing variants will be found at these sites. However, as already highlighted by the existing coding sequence variants, the definitive assignment of pathogenic significance may prove challenging. Furthermore, recent evidence from a large genome-wide meta-analysis suggests that several loci potentially involved in VWF biosynthesis and clearance may act as modifiers of the plasma VWF level. Finally, other mild bleeding disorders may be the primary cause of symptoms or may coexist in these patients with low plasma VWF levels.

**Mutational spectrum of rare bleeding disorders**

The molecular diagnosis of RBDs is based on identification of the mutation in the genes encoding the corresponding coagulation factors. Exceptions to this principle are the combined deficiency of coagulation FV + FVIII, caused by mutations in genes encoding proteins involved in the intracellular transport of FV and FVIII (MCFD2 and LMAN1) and the combined deficiency of the procoagulant vitamin K–dependent proteins (FII, FVII, FIX, and FX), caused by mutations in genes that encode enzymes involved in posttranslational modifications and in vitamin K metabolism (GGCX and VKORC1). Causative mutations in RBDs can be classified into 2 main classes: mutations abolishing the protein production or altering the secretion pathway (type I deficiency, with no antigen or activity detected in plasma) and mutations causing the production of an abnormal protein that could be completely or partially secreted but usually not active (type II deficiency with normal or reduced antigen and markedly reduced activity detected in plasma).

The 2 available registries containing data on gene mutations causing RBDs are (1) the International Registry of Rare Bleeding Disorders in Milan (www.rbdd.org), reporting information on 463 patients; and (2) the International Society on Thrombosis and Haemostasis website (http://www.isth.org/?MutationsRareBleedin), collecting information on all previously published cases. The information in the 2 registries is very similar. The detailed data on gene mutations for each gene encoding the corresponding protein are summarized in Table 2.

Missense mutations are the most frequent gene abnormalities, representing 50% to 80% of all identified mutations, except for LMAN1 variants, where the most frequent mutations are insertions/deletions (50%) and where missense mutations account for only 9% of changes. In contrast, insertion/deletion mutations represent 20% to 30% of the gene variations of the fibrinogen, FV, MCFD2, and FXIII genes and <15% of the remaining coagulation factor gene mutations. Splicing and nonsense mutations comprise 5% to 15% of
all identified mutations in all coagulation factors, with a maximum rate of 20% in the LMAN1 gene. The most infrequent types of rare bleeding disorder mutation (less than 5%) are variants located in the 3′ and 5′ untranslated regions (UTRs) of the genes, found only at the fibrinogen, FVII, FXI and FXIII loci. It should be cautioned that the functional pathogenic significance of these variants has not been established in most instances.

Although there is now a rapidly growing compendium of genetic variances that are responsible for the rare bleeding disorders, much of the early work in this area, and especially studies that have highlighted potential genotype/phenotype correlations, has been performed in mouse models. The results of the knockout animal studies in mice for each coagulation factor are reported in Table 3. The results of these studies showed that the complete absence of almost all coagulation factors involved in RBDs led to a severe clinical phenotype often incompatible with life or with reaching adulthood, except for FXI and FXIII deficiencies. These results are similar to what has subsequently been observed in humans carrying homozygous null mutations with no detectable plasma factor levels and severe bleeding symptoms.36,37 Often, bleeding manifestations occur at birth, and patients require very early prophylactic treatment.

Disruption of the fibrinogen (Fga and Fgg), Fii, Fvii, or Fx genes leads to fatal bleeding in mice.39,50 Patients with afibrinogenemia, or severe FII, FVII, and FX deficiencies, show similar bleeding manifestations and usually require regular replacement therapy. Lman1 gene knockout mice duplicate the FV- and FX-deficient phenotype in humans, albeit with a milder presentation, due to a lesser reduction in plasma FV and FVIII.44 The partial perinatal lethality observed in Lman1−/− mice on some genetic backgrounds was unexpected and has been explained as the result of a further drop in the level of LMAN1-dependent protein(s) below a critical threshold, or because of a strain-specific difference in another cargo receptor whose function overlaps with LMAN1.44

The Fxi knockout mouse model showed only a slightly increased bleeding tendency,51,52 similar to patients carrying homozygous or double-heterozygous Fxi gene mutations who usually have mild bleeding manifestations.56,58 Moreover, the results of the EN-RBD project showed that the levels of circulating FXI in plasma are not usefully associated with the clinical bleeding severity in patients.58

Mice in which the Fv gene is disrupted showed a partial embryonic lethality and a failure to reach adulthood.33 These results confirm the clinical phenotype in humans with severe FV deficiency, in whom life-threatening bleeding episodes are very rare and the main clinical features are mucosal tract bleeding56,37; however, there are some patients with FV deficiency who develop central nervous system bleeding immediately after birth and who require regular replacement therapy with fresh frozen plasma.

Fxii knockout mouse models show bleeding symptoms only in males at a later stage of life,53,56 whereas humans with severe FXIII deficiency suffer from severe bleeding symptoms from birth, with intracranial and umbilical cord bleeding being frequently documented.36,37 Female Fxiii knockout mice show only intruterine bleeding during pregnancy,53,56 similar to women with severe FXIII-A subunit deficiency who experience the same symptom, as well as recurrent abortions. Patients with a complete FXIII-B subunit deficiency show mild bleeding symptoms and normal bleeding times, whereas the mouse knockout for the Fxii-b gene shows prolonged bleeding times.56 The cause of this interspecies difference in phenotype is unknown.

**Table 2. Mutations causing RBDs**

<table>
<thead>
<tr>
<th>Deficiency</th>
<th>Gene</th>
<th>Missense</th>
<th>Nonsense</th>
<th>Splicing</th>
<th>5′-3′ UTR</th>
<th>Deletion/Insertion</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fibrinogen</td>
<td>FGA</td>
<td>39 FGA</td>
<td>18 FGA</td>
<td>7 FGA</td>
<td>3 FGA</td>
<td>38 FGA</td>
<td>145</td>
</tr>
<tr>
<td></td>
<td>FGB</td>
<td>39 FGB</td>
<td>7 FGB</td>
<td>5 FGB</td>
<td>3 FGB</td>
<td>8 FGB</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>FGG</td>
<td>67 FGG</td>
<td>4 FGG</td>
<td>7 FGG</td>
<td>1 FGG</td>
<td>253</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total = 145 (57.3%)</td>
<td>Total = 29 (11.5%)</td>
<td>Total = 19 (7.5%)</td>
<td>Total = 7 (2.8%)</td>
<td>Total = 53 (20.9%)</td>
<td></td>
</tr>
<tr>
<td>Prothrombin</td>
<td>F2</td>
<td>42 (77.8%)</td>
<td>3 (5.5%)</td>
<td>2 (3.7%)</td>
<td>0</td>
<td>7 (13%)</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>F5</td>
<td>64 (48.5%)</td>
<td>17 (12.9%)</td>
<td>15 (11.3%)</td>
<td>0</td>
<td>36 (27.3%)</td>
<td>132</td>
</tr>
<tr>
<td>Factor V</td>
<td>F7</td>
<td>151 (62.2%)</td>
<td>19 (7.8%)</td>
<td>28 (11.5%)</td>
<td>15 (6.2%)</td>
<td>30 (12.3%)</td>
<td>243</td>
</tr>
<tr>
<td>Factor VII</td>
<td>F10</td>
<td>84 (80%)</td>
<td>2 (1.9%)</td>
<td>8 (7.6%)</td>
<td>0</td>
<td>11 (10.5%)</td>
<td>105</td>
</tr>
<tr>
<td>Factor XI</td>
<td>F11</td>
<td>154 (70%)</td>
<td>23 (10.5%)</td>
<td>18 (8.2%)</td>
<td>2 (1%)</td>
<td>23 (10.5%)</td>
<td>220</td>
</tr>
<tr>
<td>Factor XIII</td>
<td>F13</td>
<td>57 (47.1%)</td>
<td>11 (9.1%)</td>
<td>17 (14%)</td>
<td>1 (1%)</td>
<td>35 (29%)</td>
<td>121</td>
</tr>
<tr>
<td>Vitamin K−dependent factors</td>
<td>GGCX</td>
<td>6 (60%)</td>
<td>0</td>
<td>3 (30%)</td>
<td>0</td>
<td>1 (10%)</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>VKORC1</td>
<td>1 (100%)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Remaining challenges in the genetic characterization of rare coagulation factor deficiencies

Despite significant advances in our knowledge of the genetic basis of the RBDs, in 5% to 10% of patients affected with severe clotting factor deficiencies, no genetic defect can be found.37

In these patients, the role of next-generation sequencing (NGS) will very likely have increasing importance. This new technology is providing enhanced opportunities for a variety of genetic strategies in which large fractions of the human genome (or even the entire genome) can be resequenced. Resequencing the exome or the entire genome is becoming the gold-standard approach for the identification of disease-causing mutations in monogenic mendelian diseases for which pathogenic variants have yet to be determined.57,59 Although the potential of these genome-wide strategies is indisputable, these approaches have yet to be used in the analysis of RBDs for which causative mutations remain elusive. By means of exome resequencing, a few unrelated, affected individuals may be sufficient to identify the causal gene for autosomal-recessive diseases, provided that the phenotype is caused by mutations in the same gene for all the individuals enrolled in the study (locus homogeneity).57,58

Finally, genotype-phenotype correlations in RBDs are difficult to establish, and the severity of bleeding is highly variable, even among patients with the same genotype. One possible explanation for the observed variability is the influence of modifier genes.
Alterations increasing risk for morbidity or death from excessive bleeding or thrombosis. Phenotypic variation can result from either rare gene variants with large effects or common gene variants with small effects, and the strategy to identify each will differ. Rare variants would include mutations in key platelet receptor genes, such as those involving the integrin αIIb gene (ITGA2B) or the integrin β3 gene (ITGB3), which give rise to Glanzmann thrombasthenia, an inherited defect of platelet cohesion (aggregation). Common variants (e.g., single-nucleotide polymorphisms or SNPs) singly are not likely to result in an overt phenotype alteration, but the cumulative or synergistic effect of multiple SNPs in key receptor genes can lead to significant differences in platelet responsiveness to various agonists or inhibitors. In general, 2 types of genetic mechanisms play a role in regulating platelet function, and these same principles also apply to the other bleeding phenotypes alluded to in the sections on hemophilia, VWD, and the RBDs.

1. Common variants with small effects

It has been accepted for several decades that variation in platelet responsiveness to 1 or more common agonists is heritable. Microarray technology enabled the analysis of the contribution of candidate genes to variation in platelet responsiveness. Generally speaking, the

---

**Inherited platelet-related bleeding disorders**

Because platelets play a central role in the maintenance of hemostasis, genotypic variations can result in significant phenotypic alterations increasing risk for morbidity or death from excessive bleeding or thrombosis. Phenotypic variation can result from either rare gene variants with large effects or common gene variants with small effects, and the strategy to identify each will differ. Rare variants would include mutations in key platelet receptor genes, such as those involving the integrin αIIb gene (ITGA2B) or the integrin β3 gene (ITGB3), which give rise to Glanzmann thrombasthenia, an inherited defect of platelet cohesion (aggregation). Common variants (e.g., single-nucleotide polymorphisms or SNPs) singly are not likely to result in an overt phenotype alteration, but the cumulative or synergistic effect of multiple SNPs in key receptor genes can lead to significant differences in platelet responsiveness to various agonists or inhibitors. In general, 2 types of genetic mechanisms play a role in regulating platelet function, and these same principles also apply to the other bleeding phenotypes alluded to in the sections on hemophilia, VWD, and the RBDs.

---

**Inherited platelet-related bleeding disorders**

Because platelets play a central role in the maintenance of hemostasis, genotypic variations can result in significant phenotypic alterations increasing risk for morbidity or death from excessive bleeding or thrombosis. Phenotypic variation can result from either rare gene variants with large effects or common gene variants with small effects, and the strategy to identify each will differ. Rare variants would include mutations in key platelet receptor genes, such as those involving the integrin αIIb gene (ITGA2B) or the integrin β3 gene (ITGB3), which give rise to Glanzmann thrombasthenia, an inherited defect of platelet cohesion (aggregation). Common variants (e.g., single-nucleotide polymorphisms or SNPs) singly are not likely to result in an overt phenotype alteration, but the cumulative or synergistic effect of multiple SNPs in key receptor genes can lead to significant differences in platelet responsiveness to various agonists or inhibitors. In general, 2 types of genetic mechanisms play a role in regulating platelet function, and these same principles also apply to the other bleeding phenotypes alluded to in the sections on hemophilia, VWD, and the RBDs.

---

**Table 3. Results of coagulation factor gene disruption in mice**

<table>
<thead>
<tr>
<th>Deficiency</th>
<th>Knockout mouse model</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fibrinogen</td>
<td>Figg: Mice homozygous for disruptions of this gene have blood that is unable to clot. On some genetic backgrounds, this can lead to fatal bleeding. Figg: Pregnant homozygous null mice exhibit retarded embryoplacental development, spontaneous abortion, and maternal death through excessive uterine bleeding. Mutants expressing a truncated polypeptide show reduced platelet aggregation, increased bleeding time, and occasional fatal neonatal bleeding.</td>
<td>39,40</td>
</tr>
<tr>
<td>Prothrombin</td>
<td>Partial embryonic lethality (E9.5-E11.5). Not compatible with survival to adulthood. Neonatal hemorrhage, no survival beyond several days, loss of vascular integrity.</td>
<td>41,42</td>
</tr>
<tr>
<td>Factor V</td>
<td>Half of mice homozygous for a null allele die at E9 to E10 with defects in yolk-sac vasculature and somite formation; the remaining half develop to term but die of massive hemorrhage within hours of birth.</td>
<td>43</td>
</tr>
<tr>
<td>Factor V and Factor VIII</td>
<td>Levels of plasma FV and FVIII are all reduced to 50% of wild-type in Lman1 KO mice, compared with the 5% to 30% levels typically observed in human F5F8D patients. Lack of a significant bleeding phenotype. Partial lethal phenotype was observed in Lman1 KO mice only on certain genetic backgrounds.</td>
<td>44</td>
</tr>
<tr>
<td>Factor VII</td>
<td>No embryonic lethality. Not compatible with survival to adulthood. Normal embryonic development, death caused by hemorrhage in neonates, no survival beyond 24 d, no vascular defects.</td>
<td>45-47</td>
</tr>
<tr>
<td>Factor X</td>
<td>Partial embryonic lethality (E11.5-E12.5). Complete absence of FX is incompatible with murine survival to adulthood but minimal FX activity as low as 1% to 3% is sufficient to rescue the lethal phenotype. Death caused by hemorrhage in neonates, no survival beyond 20 d, no vascular defects.</td>
<td>48-50</td>
</tr>
<tr>
<td>Factor XI</td>
<td>Mice homozygous for a knockout allele show a tendency for slightly prolonged tail transection bleeding times and are protected from vessel-occluding fibrin formation after transient ischemic brain injury (<a href="http://www.informatics.jax.org/">http://www.informatics.jax.org/</a>).</td>
<td>51,52</td>
</tr>
<tr>
<td>Factor XIII-A</td>
<td>No embryonic lethality. Compatible with survival to adulthood. However, the survival rate of Fxiii-A KO males at 10 mo was markedly lower than that of the wild-type males. In dead mice, intrathoracic hemorrhage and large hematoma were found. The survival rate of Fxiii-A KO females did not differ. Heart severe fibrosis together with hemosiderin deposition in Fxiii-A KO males. Normal embryonic, all homozygous KO mice appeared normal at birth, about half of the pregnant KO females died due to excessive vaginal bleeding.</td>
<td>53-55</td>
</tr>
<tr>
<td>Factor XIII-B</td>
<td>Significantly reduced level of FXIII A subunit in plasma. Prolonged bleeding time. Compatible with survival to adulthood. No Fxiii-B KO female mice died during pregnancy, and these animals were capable of successfully carrying multiple litters. A low degree of fibrosis in the myocardium was also seen in Fxiii-B KO males.</td>
<td>55,56</td>
</tr>
</tbody>
</table>

E, embryonic day.
results of these studies failed to produce results that were consistently replicated in different populations in even well-defined large population cohorts by because of (1) lack of a single reproducible phenotype that is representative of platelet function; (2) the use of inadequately sized, population isolates; and (3) an incomplete knowledge of common alleles or haplotype structure in the available human populations.  

Despite these weaknesses, there were several candidate gene polymorphisms that generated consistent associations, including those located in GP6, ITGA2, P2Y12, ADRA2A, and CYP2C19, subsequently validated by functional genomic or genome-wide association studies. Biological validation of certain SNPs, such as GP6 (rs1671152), was particularly impressive. 

Genome-wide association studies have also identified several common SNPs, most in noncoding regions, that correlate with platelet function, mean platelet volume, or platelet count. Roughly one-third of these SNPs are located close to genes that can be associated with megakaryocyte/platelet function,65,66 mean platelet volume, or platelet count.  

This information can assist basic researchers in clarifying the role of individual genes and gene networks in megakaryocyte/platelet physiology, but knowledge of common sequence variants is currently of little clinical use because the effect sizes on the risk for platelet dysfunction or adverse outcomes are small. However, the increasing use of NGS is beginning to change this situation by facilitating the generation of a rapidly accumulating database of these variants. 

The development of a global database using NGS is being driven by cooperative efforts to decipher the exomes or complete genomes of hundreds of thousands of donors of defined racial/ethnic backgrounds in order to establish a useful catalog of human sequence variants. 

Some of these variants will very likely become part of the routine workup of patients, particularly those with early-onset defects of hemostasis or platelet function.  

2. Rare variants with large effects

A number of platelet function disorders (PFDs) have been extensively characterized and shown to be paucigenic, if not monogenic (Table 4). In certain cases, the causative rare gene variants are often identified and readily validated. However, in other PFDs, clinical and laboratory workups usually do not provide sufficient evidence to target the candidate gene(s). Enrichment of target DNA together with NGS will facilitate diagnosis by enabling a molecular analysis of large numbers of candidate genes. The first step is the generation of a candidate gene list, permitting NGS of a limited and relevant gene set as part of early diagnosis. One example of such a PFD candidate gene set has been developed by the UK Genotyping and Phenotyping study group.

Genetic characterization of inherited platelet disorders

The text below highlights several examples of NGS successes in the identification of loci responsible for monogenic inherited thrombocytopenias and PFDs. Until relatively recently, the genetic characterization of inherited platelet disorders had been confined to the more prevalent (but nevertheless, still rare) conditions such as Glanzmann thrombasthenia, Bernard Soulier syndrome, and platelet-type VWD. Genetic characterization of the conditions detailed below illustrates the enhanced potential for pathogenetic understanding with NGS strategies.

Familial platelet disorder with predisposition to acute myeloid leukemia

Familial platelet disorder with predisposition to acute myeloid leukemia (AML) is an autosomal-dominant disorder characterized with megakaryocyte/platelet function. However, the increasing use of NGS is beginning to change this situation by facilitating the generation of a rapidly accumulating database of these variants. 

The development of a global database using NGS is being driven by cooperative efforts to decipher the exomes or complete genomes of hundreds of thousands of donors of defined racial/ethnic backgrounds in order to establish a useful catalog of human sequence variants. Some of these variants will very likely become part of the routine workup of patients, particularly those with early-onset defects of hemostasis or platelet function.

<table>
<thead>
<tr>
<th>Disease</th>
<th>Gene(s)</th>
<th>Affected protein(s)</th>
<th>Megakaryocyte/platelet function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glanzmann thrombasthenia</td>
<td>ITGA2B, ITGB3</td>
<td>Integrin αIIbβ3</td>
<td>Binding to fibrinogen or VWF</td>
<td>75,81</td>
</tr>
<tr>
<td>Bernard-Soulier syndrome</td>
<td>GP1BA, GP1BB, GP9</td>
<td>Glycoprotein Ibα</td>
<td>Binding to VWF or thrombin</td>
<td>82</td>
</tr>
<tr>
<td>Gray platelet syndrome</td>
<td>NBEAL2</td>
<td>Neurobeachin-like protein 2</td>
<td>Vesicular trafficking; platelet α granule development</td>
<td>88-90</td>
</tr>
<tr>
<td>Familial platelet disorder with a predisposition to AML</td>
<td>RUNX1</td>
<td>Runt-related transcription factor 1</td>
<td>Transcription factor involved in megakaryocyte differentiation</td>
<td>86,87</td>
</tr>
<tr>
<td>Thrombocytopenia 2</td>
<td>ANKRD26</td>
<td>Ankyrin repeat domain-containing protein 26</td>
<td>Unknown</td>
<td>91,92</td>
</tr>
<tr>
<td>Thrombocytopenia-absent radius syndrome</td>
<td>RBM6A</td>
<td>RNA-binding protein 8A</td>
<td>Unknown</td>
<td>93</td>
</tr>
<tr>
<td>Wiskott-Aldrich syndrome</td>
<td>WAS</td>
<td>Wiskott-Aldrich syndrome protein</td>
<td>Actin cytoskeleton organization and signaling</td>
<td>74</td>
</tr>
<tr>
<td>Chediak-Higashi syndrome</td>
<td>LYST</td>
<td>Lysosomal-trafficking regulator</td>
<td>Lysosomal trafficking</td>
<td>77</td>
</tr>
<tr>
<td>MYH9-related disease</td>
<td>MYH9</td>
<td>Myosin heavy chain 9</td>
<td>Proplatelet formation</td>
<td>77,78</td>
</tr>
<tr>
<td>Hermansky-Pudlak syndrome</td>
<td>AP3B1, BLOC1S3</td>
<td>AP-3 complex subunit β-1</td>
<td>Biogenesis of lysosomal-related organelles complex</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>BLOC1S6, DTNBPI, HPS1, HPS3, HPS4, HPS5, HPS6, HPS7, HPS8, HPS9</td>
<td></td>
<td>Formation of lysosomal-related organelles</td>
<td></td>
</tr>
</tbody>
</table>

The text below highlights several examples of NGS successes in the identification of loci responsible for monogenic inherited thrombocytopenias and PFDs. Until relatively recently, the genetic characterization of inherited platelet disorders had been confined to the more prevalent (but nevertheless, still rare) conditions such as Glanzmann thrombasthenia, Bernard Soulier syndrome, and platelet-type VWD. Genetic characterization of the conditions detailed below illustrates the enhanced potential for pathogenetic understanding with NGS strategies.

Familial platelet disorder with predisposition to acute myeloid leukemia

Familial platelet disorder with predisposition to acute myeloid leukemia (AML) is an autosomal-dominant disorder characterized
by moderate thrombocytopenia, dysmegakaryopoiesis, variable platelet defects, and a propensity to develop myelodysplastic syndromes and/or AML. Germline mutations or deletions in RUNX1 are thought to be causative.

Hermansky-Pudlak syndrome (HPS)

The rapid diagnosis of a HPS in a single test was enabled by NGS of a set of 57 genes implicated in secretion, leading to the identification of a causative single variant of the gene HPS4 in 1 patient.

Gray platelet syndrome

Exome and RNA sequencing of twenty patients by three independent groups led to the discovery that NBEAL2 is the causative gene for gray platelet syndrome.

Thrombocytopenia 2, also known as ankyrin repeat domain 26 (ANKRD26)-related thrombocytopenia (ANKRD26-RT)

Thrombocytopenia 2 may account for 10% or more of all inherited thrombocytopenias. Twelve different mutations in a very limited region (22 bp) of the 5′ UTR of ANKRD26 have been identified in 21 pedigrees, result in enhanced transcriptional activity of ANKRD26, and are considered causative. The diagnosis can be made by mutation screening of the relevant 5′-UTR sequence (~527 bp).

Thrombocytopenia and absent radii syndrome

Through exome sequencing of pedigrees of 53 cases, it was determined that the compound inheritance of a low-frequency regulatory SNP and a rare null mutation in the RBM8A gene causes thrombocytopenia and absent radii syndrome.

Wiskott-Aldrich syndrome (WAS)

WAS results from mutations in the WAS gene, encoding WASP, which plays a critical role in actin cytoskeleton organization and signaling. In 77 WAS families, gene sequencing identified 62 mutations including 17 novel sequence variants.

Future genomic strategies for the characterization of inherited bleeding disorders

As genomic technologies continue to advance, the potential for incorporating these strategies into diagnostic algorithms will inevitably increase. However, before this can happen, a number of critical additions to our current diagnostic infrastructure must be developed. Carefully curated databases, such as Reactome, provide key information that will enable physicians and scientists to access relevant information on cellular pathways and to visualize the relationships between rare sequence variants and clinical phenotypes. The GEN2PHEN database, created through a collaboration between the National Center for Bioinformatics and European Bioinformatics Institute, and other efforts have resulted in the establishment of the Locus Reference Genome (www.lrg-sequence.org), a system that provides a genomic DNA sequence representing a single gene, whose core content never changes, with an updateable annotation layer that includes sequence variants.

Working exclusively in the area of hemostasis/thrombosis, the ThromboGenomics project (http://haemgen.haem.cam.ac.uk/), sponsored by the International Society on Thrombosis and Haemostasis, was recently established to facilitate the introduction of genomic approaches to the diagnosis of inherited bleeding disorders. There are several goals for this initiative: to provide gene annotation relevant to clinical standards, to develop a process of database curation that ensures a stable and sustainable frame of reference, and to test the proof-of-principle that NGS can improve the diagnosis of inherited bleeding disorders. To date, the curation for 4 platelet disorders (Glanzmann thrombasthenia, Bernard-Soulier syndrome, Wiskott-Aldrich syndrome and familial platelet disorder with predisposition to acute myelogenous leukemia) has been completed, and locus reference genomes have been created for 7 genes: ITGA2B, ITGB3, GP1BA, GP1BB, GP9, WAS, and RUNX1. Functional variants have been agreed for these genes, and a first-generation sequencing platform is currently in development. The ThromboGenomics initiative is also planning to develop a custom sequence panel, capturing nearly 1 million base pairs including all regions containing bleeding and platelet disorder variants present in the Human Gene Mutation Database for 96 candidate genes involved in hemostasis.

Acknowledgments

D.L. is the recipient of a Canada Research Chair in Molecular Hemostasis. His studies of hemophilia and VWD are supported by the Canadian Institutes of Health Research. F.P. is the recipient of the 2011 Bayer Hemophilia Award – Special Project Award and is the leader of the work package 8 of the project “European Haemophilia Network (EUHANET)” funded by the European Commission via the Executive Agency for Health and Consumers (EAHC) (agreement number 20111207).

Authorship

Contribution: F.P., T.K., and D.L. all contributed to the design and writing of this review article.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: David Lillicrap, Department of Pathology and Molecular Medicine, Richardson Laboratory, Queen’s University, Kingston, Ontario K7L 3N6, Canada; e-mail: dpl@queensu.ca.

References

6. Keene S, Mitchell M, Goodeve A; UK Haemophilia Centers Doctors’ Organization Haemophilia Genetics Laboratory Network. The


Genetic sequence analysis of inherited bleeding diseases

Flora Peyvandi, Tom Kunicki and David Lillicrap