A sensitized mutagenesis screen in Factor V Leiden mice identifies novel thrombosis suppressor loci

Randal J. Westrick1,2,3,*, Kärt Tomberg3,4,*, Amy E. Siebert1, Guojing Zhu3, Mary E. Winn3, Sarah L. Dobies3, Sara L. Manning3, Marisa A. Brake1, Audrey Cleuren3, Lena M. Mishack1, Linzi M. Hobbs1, Alexander Johnston1, David R. Siemieniak5, Jishu Xu4, Jun Z. Li4 and David Ginsburg3,4,5,6

1Oakland University Department of Biological Sciences
2Oakland University Center for Data Science and Big Data Analysis
3Life Sciences Institute, University of Michigan
4Department of Human Genetics, University of Michigan
5Howard Hughes Medical Institute
6Departments of Internal Medicine and Pediatrics, University of Michigan

*These authors contributed equally to this work

Co-corresponding authors:

David Ginsburg, MD. 5214 LSI Building, 210 Washtenaw Avenue, Ann Arbor MI 48109. E-mail: ginsburg@umich.edu, Telephone: 734-647-4808, Fax: 734-936-2888.

Randal Westrick, PhD. 1800 Library Drive, 305 Dodge Hall, Rochester MI, 48309-4479 E-mail: rjwestrick@oakland.edu, Telephone: 248-370-3577.
Abstract

Factor V Leiden (F\textsuperscript{5L}) is a common genetic risk factor for venous thromboembolism (VTE), though it exhibits only 10% penetrance. We conducted a sensitized ENU mutagenesis screen for dominant thrombosis-suppressor genes based on perinatal lethal thrombosis in mice homozygous for F\textsuperscript{5L} (F\textsuperscript{5L/L}) and haploinsufficient for tissue factor pathway inhibitor (T\textit{fpi}\textsuperscript{+/−}). The observation that F8 deficiency enhanced survival of F\textsuperscript{5L/L} T\textit{fpi}\textsuperscript{+/−} mice demonstrated the potential for genetic suppression of F\textsuperscript{5L/L} T\textit{fpi}\textsuperscript{+/−} lethality. G\textsubscript{0} ENU-mutagenized F\textsuperscript{5L} males and F\textsuperscript{5L/−} T\textit{fpi}\textsuperscript{+/−} females were next crossed to generate 6,739 G1 progeny, with 98 F\textsuperscript{5L/L} T\textit{fpi}\textsuperscript{+/−} offspring surviving until weaning and 16 exhibiting transmission of a putative thrombosis-suppressor to subsequent generations. The resulting lines are referred to as MF5L, (M\textit{odifier of F}actor 5 L\textit{eiden 1-16}). Linkage analysis in MF5L6 identified a chromosome 3 locus containing the tissue factor gene (F3). Though no ENU-induced F3 mutation was identified, heterozygous F3 deficient mice (F3\textsuperscript{+/−}) suppressed F\textsuperscript{5L/L} T\textit{fpi}\textsuperscript{+/−} lethality. Thus, like F8 deficiency, reduced F3 activity suppresses F\textsuperscript{5L/L} T\textit{fpi}\textsuperscript{+/−} thrombosis. Whole exome sequencing in MF5L12 identified an Actr2 gene point mutation (p.R258G) as the sole candidate. Attempts to generate an independent Actr2 knockin/knockout via CRISPR/Cas9 failed for any mouse beyond the 60-cell stage. Our findings identify F8 and the TFPI/F3 axis as key regulators of thrombosis balance in the setting of F\textsuperscript{5L} and demonstrate the utility of this sensitized ENU mutagenesis approach for the identification of dominant thrombosis suppressor loci.
Introduction

Venous thromboembolism (VTE) is a common disease that affects 1 to 3 per 1000 individuals per year\(^1\). VTE susceptibility exhibits a complex etiology involving contributions of both genes and environment. Genetic risk factors explain approximately 60% of the overall risk for VTE\(^2\). Recent large-scale genome wide association studies (GWAS) confirm \(ABO\), \(F5\), \(FGG\) and \(F2\) as thrombosis susceptibility genes, with few additional novel loci identified\(^3\)\(^-\)\(^5\), leaving the major component of VTE genetic risk still unexplained.

The Factor V Leiden variant (\(F5^L\)) is a common inherited risk factor for VTE with an allele frequency of 2-10% in most European-derived populations\(^6\),\(^7\). \(F5^L\) is estimated to account for up to 25% of the genetically-attributable thrombosis risk in humans\(^6\). However, penetrance is incomplete, with only 10% of \(F5^L\) heterozygotes developing thrombosis in their lifetimes. The severity of thrombosis also varies widely among affected individuals\(^7\),\(^8\), limiting the clinical utility of \(F5^L\) genotyping in the management of VTE\(^9\).

The incomplete penetrance and variable expressivity of thrombosis among \(F5^L\) patients can at least partially be explained by genetic interactions between \(F5^L\) and other known thrombotic risk factors such as hemizygosity for antithrombin III or proteins C or S, as well as the common prothrombin 20210 polymorphism\(^8\),\(^10\),\(^11\). However, <2% of \(F5^L\) heterozygotes would be expected to co-inherit a mutation at one or more of these loci, suggesting that a large number of additional genetic risk factors for VTE and/or modifiers of \(F5^L\) remain to be identified\(^3\),\(^8\).

Mice carrying the orthologous \(F5^L\) mutation exhibit a mild to moderate prothrombotic phenotype closely mimicking the human disorder\(^12\). We previously reported a synthetic lethal interaction between \(F5^L\) homozygosity (\(F5^{L/L}\)) and hemizygosity for tissue factor pathway
inhibitor (Tfpi<sup>+/−</sup>)<sup>13</sup>. Nearly all mice with this lethal genotype combination (F5<sup>L/L</sup> Tfpi<sup>+/−</sup>) succumb to widespread, systemic thrombosis in the immediate perinatal period<sup>13</sup>.

ENU mutagenesis in mice has been used effectively to identify novel genes involved in a number of biological processes<sup>14</sup>,<sup>15</sup>. ENU-induced germline mutations transmitted from a mutagenized male mouse (G0) occur at ~1.5 mutations per megabase, at least 50 fold higher than the endogenous background mutation rate<sup>16</sup>. Several previous reports have successfully applied an existing phenotype as a sensitizer to identify modifier genes. A dominant suppressor screen in MecP2 deficient mice (Rett syndrome) identified a mutation in squalene epoxidase (Sqle) as a heritable suppressor, resulting in prolonged survival and amelioration of neurologic manifestations<sup>17</sup>. Other successful sensitized screens include analysis of mouse mutants predisposed to diabetic nephropathy<sup>18</sup>, a screen in Sox10 haploinsufficient mice identifying the Gli3 gene as a modifier of neurochristopathy<sup>19</sup> and identification of a mutation in the c-Myb gene as a dominant modifier for platelet count in Mpl deficient mice (congenital thrombocytopenia)<sup>20</sup>.

We now report the results of a dominant, sensitized ENU mutagenesis screen for suppressors of F5<sup>L/L</sup> Tfpi<sup>+/−</sup> dependent lethal thrombosis.

Results and Discussion

F8 deficiency suppresses F5<sup>L/L</sup> Tfpi<sup>+/−</sup> lethality

To test whether the F5<sup>L/L</sup> Tfpi<sup>+/−</sup> lethal phenotype was genetically suppressible by F8 deficiency (X-linked hemophilia A in humans), triple heterozygous F5<sup>L/+</sup> Tfpi<sup>+/−</sup> F8<sup>+/−</sup> female mice were generated and crossed to F5<sup>L/L</sup> male mice (Figure 1A). One quarter of conceptuses are expected to carry the F5<sup>L/L</sup> Tfpi<sup>+/−</sup> genotype, with half of the total expected male conceptuses completely F8 deficient (F8<sup>−</sup>), and thus 1/16<sup>th</sup> of the overall offspring from this mating expected
to be $F5^{L/L}$ $Tfpi^{+/}$ $F8^{-/-}$ (males). Similarly, $1/16^{th}$ of the progeny should be $F5^{L/L}$ $Tfpi^{+/}$ $F8^{-/-}$ (females). A total of 167 progeny from this cross were genotyped at weaning, with 8 $F5^{L/L}$ $Tfpi^{+/}$ $F8^{-/-}$ male mice observed (compared to 0 $F5^{L/L}$ $Tfpi^{+/}$ $F8^{+/-}$, $p=0.01$) and 2 $F5^{L/L}$ $Tfpi^{+/}$ $F8^{+/-}$ female mice (compared to 1 $F5^{L/L}$ $Tfpi^{+/-}$ $F8^{+/-}$, N.S.) (Sup. Table 1). These data strongly suggest that the $F5^{L/L}$ $Tfpi^{+/}$ phenotype is genetically suppressible and are consistent with human studies demonstrating $F8$ level as an important VTE risk factor$^{21}$.

**The $F5^{L/L}$ $Tfpi^{+/}$ phenotype is suppressed by dominant ENU induced mutations**

A sensitized, genome-wide ENU mutagenesis screen for dominant thrombosis suppressor genes was implemented as depicted in Figure 1B. ENU mutagenized G0 $F5^{L/L}$ males were crossed to $F5^{L/+}$ $Tfpi^{+/-}$ females to generate G1 mice, which were screened by genotyping at weaning for $F5^{L}$ and $Tfpi^{+/-}$. Previously described visible dominant mutants$^{22}$, including belly spotting and skeletal abnormalities, were observed in approximately 5.9% of G1 offspring, similar to the ~4.2% rate of observable mutants in previous studies$^{22}$, and consistent with the ~20-30 functionally significant mutations per G1 mouse expected with this ENU mutagenesis protocol$^{23}$, $^{24}$. One quarter of G1 embryos from this cross would be expected to carry the synthetic lethal $F5^{L/L}$ $Tfpi^{+/-}$ genotype. Out of a total of 6,739 G1 mice screened at weaning, the 98 live $F5^{L/L}$ $Tfpi^{+/-}$ mice (45 females, 53 males, Sup. Table 2) represented 4.4% of the 2,214 embryos expected with this genotype.

The heritability of each of the 98 G1 putative suppressor mutants was evaluated by a progeny test backcross to C57BL/6J (B6) $F5^{L/L}$ mice. The observation of one or more $F5^{L/L}$ $Tfpi^{+/-}$ offspring surviving to weaning increased the likelihood that a particular modifier of Factor $5$ Leiden ($MF5L$) line carries a transmissible suppressor mutation. 75 of the 98 surviving $F5^{L/L}$ $Tfpi^{+/-}$ G1 mice produced no offspring surviving to weaning, either due to early lethality or
infertility, with over 1/2 of these mice (37 of 75) exhibiting a grossly runted appearance. Survival data were available for 57 of the $F5^{L/L}$ $Tfp^{+/-}$ G1 mice, with 34 living past 70 days of age with a trend toward poorer survival for females (Figure 1C). Of the 23 $F5^{L/L}$ $Tfp^{+/-}$ G1 mice producing 1 or more G2 progeny surviving to weaning, 7 produced no $F5^{L/L}$ $Tfp^{+/-}$ G2s, including 4 G1s with 8 or more offspring of other genotypes. Sixteen $F5^{L/L}$ $Tfp^{+/-}$ G1 mice (4 female, 12 male) produced one or more $F5^{L/L}$ $Tfp^{+/-}$ progeny when bred to B6 $F5^{L/L}$ mice (Sup Table 3). These 16 putative “rescue” mice were crossed into 129S1/SvIMJ (129S1) to generate suppressor lines of genetically informative progeny for genetic mapping. The number of total progeny, genotypic distribution and penetrance of the $F5^{L/L}$ $Tfp^{+/-}$ mice in each line are listed in Sup. Table 3. Within these suppressor lines, mice with the $F5^{L/L}$ $Tfp^{+/-}$ genotype were ~30% smaller than their $F5^{L/L}$ littermates at the time of weaning ($p<2.2\times10^{-16}$, Figure 1D), with this difference maintained after outcrossing to the 129S1 strain (Figure 1E).

Previous reports based on the specific locus test estimate an ENU-induced mutation rate of 1/700 visible mutations per locus for the ENU dosing regimen used here$^{28}$. This mutation rate predicts that our screen of 6,739 G1 progeny ($2214 F5^{L/L} Tfp^{+/-}$ expected) should have resulted in ~3 mutations per gene averaged over the entire genome, with 54% of these mutations expected to be null based on gene function in a specific locus test (28), for an overall estimated genome coverage with null mutations of ~1.5X.

**The $MF5L6$ suppressor mutation maps to a chromosome 3 interval containing $F3$**

In order to map putative ENU-induced suppressor mutations, surviving $F5^{L/L}$ $Tfp^{+/-}$ mice were intercrossed with $F5^{L/L}$ mice that had been extensively backcrossed onto the 129S1 strain. Crosses between $F5^{L/L}$ and $F5^{L/+} Tfp^{+/-}$ mice (both $F5^{L}$ and $Tfp^{+/-}$ backcrossed > 12 generations
onto 129S1) confirmed the lethality of the $F5^{+/L}\ Tfpi^{+/-}$ genotype on the 129S1 background (Sup. Table 4).

Each $MF5L$ suppressor line was intercrossed to 129S1 mice, with the 4 largest lines ($MF5L1, 6, 9$ and $16$) producing greater than 13 surviving $F5^{L/L}\ Tfpi^{+/-}$ offspring on the mixed 129S1-B6 genetic background. Though the $MF5L1, MF5L9$ and $MF5L16$ were successfully expanded to pedigrees containing 27, 84, and 14 $F5^{L/L}\ Tfpi^{+/-}$ informative offspring, respectively, genotyping for a total of ~800 markers in each cross failed to identify any loci with a LOD $\geq 3$ (Sup. Table 5). The absence of a clear linkage signal for each of these lines likely reflects complex mouse strain modifier gene interactions, which are known to significantly impact mouse phenotypes$^{8,25}$ and confound linkage analysis$^{26}$. Consistent with this hypothesis, analysis of $F5^{L/L}$ and $F5^{L/+}$ littermates demonstrated consistently poorer survival in the purebred 129S1 versus B6 strain backgrounds$^{12}$ (Sup. Table 4, Sup. Table 6).

$MF5L6$ was maintained for 12 generations on both the mixed and B6 backgrounds and produced a total of 336 $F5^{L/L}\ Tfpi^{+/-}$ mice (98 of which were genetically heterogeneous and therefore useful for linkage analysis). Genome-wide SNP genotyping was performed on DNA from 98 genetically informative $F5^{L/L}\ Tfpi^{+/-}$ mice, with multipoint linkage analysis shown in Figure 2A. Since the genetic intervals around the $F5$ and $Tfpi$ loci cannot be accurately assessed for linkage, these regions of chromosomes 1 and 2 were excluded from linkage analysis (See figure legend or Methods). A single locus with a maximum LOD score $>3.3$ was identified on Chr 3 (maximum LOD=4.49), with the 1 LOD interval (117.3-124.8Mb) containing 38 refseq annotated genes (Figure 2C).

The $F3$ gene located at Chr3:121.7 Mb (Figure 2C) encodes Tissue factor (TF), a procoagulant component of the hemostatic pathway that has $Tfpi$ as its major regulator.
Quantitative or qualitative deficiencies in F3 are thus highly plausible candidates to suppress the $F5^{L/L}$ Tfpi$^{+/-}$ phenotype. However, analysis of sequencing data from the full set of F3 exons and introns as well as the 5kb upstream of exon 1, failed to identify an ENU-induced mutation. Analysis of F3 mRNA levels in liver, lung and brain tissues of adult mice failed to identify any differences in the level of expression from the ENU-mutant compared to the wildtype allele (Sup. Fig 1). However, an ENU-induced regulatory mutation outside the sequenced region that confers a key change in tissue- or developmental stage-specific F3 mRNA expression, or in response to environmental stress, cannot be excluded.

Whole exome sequencing data analysis of a $F5^{L/L}$ Tfpi$^{+/-}$ mouse from MF5L6 (see below) failed to identify an ENU variant in F3 or the other genes on chromosome 3. Although additional ENU variants were identified on other chromosomes, none co-segregated with the survival phenotype in line MF5L6 (Sup. Table 8).

**F3 haploinsufficiency suppresses the $F5^{L/L}$ Tfpi$^{+/-}$ lethal phenotype**

To test F3 as a candidate suppressor of the $F5^{L/L}$ Tfpi$^{+/-}$ phenotype, an independent F3 null allele was introduced and triply heterozygous $F5^{L/+}$ Tfpi$^{+/-}$ F3$^{+/-}$ mice crossed to $F5^{L/L}$ B6 mice (Figure 2B). Of 272 progeny genotyped at weaning, 13 $F5^{L/L}$ Tfpi$^{+/-}$ F3$^{+/-}$ were observed (compared to 1 $F5^{L/L}$ Tfpi$^{+/-}$ F3$^{+/-}$, p=0.004, Sup. Table 7). We also observed significantly fewer male than female $F5^{L/L}$ Tfpi$^{+/-}$ F3$^{+/-}$ mice (2 vs. 11 p=0.03). Thus, haploinsufficiency for F3$^{+/-}$ suppresses the synthetic lethal $F5^{L/L}$ Tfpi$^{+/-}$ phenotype, though with incomplete penetrance (33%) that also differs by gender (10% for females and 67% for males). Gender specific differences in venous thrombosis rates have previously been reported, including contributions from oral contraceptives and hormone replacement therapy$^{27, 28}$. Conversely, the MF5L6 line had an overall penetrance of 72.4%, with similar male/female penetrance. This difference in penetrance
could be due to 129S1 strain effects in the MF5L6 line. Our Sanger sequencing data of F3 exons and regulatory regions identified no mutations. However, enhancers and repressors greater than 1 megabase away from a gene’s transcription start site have been documented to control gene expression. Taken together, these data suggest that an ENU-induced F3 regulatory mutation outside of the sequenced segment may be responsible for thrombosuppression in MF5L6, though we cannot exclude regulatory mutation in another gene. Nonetheless, our findings demonstrate that F3/Tfpi balance plays a key role in VTE in the mouse, particularly in the setting of F5L, and suggest that modest variations in either F3 or Tfpi could be important modifiers of VTE susceptibility in humans.

**Whole exome sequencing identifies candidate ENU suppressor variants for 8 MF5L lines**

Whole exome-next generation sequencing (NGS) was performed on genomic DNA from one F5/L Tfpi+/− mouse of varying generations (ranging from G2 to G5), from each of 8 MF5L lines, including the 4 lines described above, as well as 4 additional lines with large pedigrees (MF5L5, MF5L8, MF5L11, MF5L12). The mean coverage of sequenced exomes was more than 90X, with >97% of the captured region covered with at least 6 independent reads (Sup. Table 8). A total of 125 heterozygous variants were identified as candidate suppressor mutations, with 79 variants occurring within coding sequence. Of these latter mutations, 54.5% were nonsynonymous single nucleotide variants (SNVs), followed by UTR (17.6%), synonymous (14.4%) and stopgain SNVs (7.2%), with the remainder being comprised of indels, splicing, and stoploss mutations. The most common mutation events were A/T→G/C transitions (35.2%), while C/G→G/C transversions were the least represented (2.5%). This spectrum of mutations is consistent with previously published ENU reports. Validation was performed for 52 variants by Sanger sequencing, including all nonsynonymous and stopgain mutations. Variants exhibiting no
recombination with the *Tfpi* locus on chromosome 2 were excluded from further analysis (See methods of Fig 2A). The remaining variants were then checked for parent of origin (either the G1 mutagenized progeny or its nonmutagenized mate) as well as the original mutagenized G0 male. 42 of the variants were identified in the G1 mouse but not in the G0 or nonmutagenized parent, consistent with an ENU-induced mutations. The remaining 10 mutations were either not verified or were transmitted from the non-mutagenized parent, indicating either spontaneous mutations in the nonmutagenized parent or errors in the sequencing data. (Sup. Table 9).

Each SNV was analyzed for validation in additional mice *MF5L* mice from the line in which it was identified. None of the thrombosuppressive exonic ENU-induced variants were validated in lines *MF5L1*, 5, 6, 8, 11 and 16 (either they did not segregate with the lethal phenotype or they were exome sequencing errors). Of the 7 candidate ENU-induced SNVs identified from whole exome sequencing analysis for the *MF5L12* line, 6 were validated by Sanger sequencing as consistent with ENU-induced mutations in the G0 mice (Sup. Table 10). For each of these 6 SNVs, co-segregation with the survival phenotype was tested by Kaplan-Meier analysis of 31 *F5L/L Tfpi+/−* mice from the *MF5L12* line. Only one variant, a nonsynonymous SNV in the *Actr2* gene (*Actr2+/G*), demonstrated a significant survival advantage\(^\text{22}\) when co-inherited with the *F5L/L Tfpi+/−* genotype (p=1.7x10\(^{-6}\)) (Figure 3A).

The *Actr2* gene encodes the ARP2 protein, which is an essential component of the ARP2/3 complex\(^\text{30}\). The *Actr2+/G* mutation results in a p.R258G substitution in exon 7 of *Actr2*, at a highly conserved amino acid position, with arginine present at this position for all 60 available vertebrate sequences ([https://genome.ucsc.edu](https://genome.ucsc.edu)), as well as in plants and fungi (Figure 3B). In addition, no variants at this position have been identified to date in over 120,000 human alleles ([ExAC, [http://exac.broadinstitute.org](http://exac.broadinstitute.org) accessed: 10/2016]). The ARP2/3 complex is
essential for actin branching and polymerization and complete ARP3 deficiency is embryonic lethal in mice\textsuperscript{31}. The other members of the complex include ARPC 1-5. Disruption of any one of the members of the ARP2/3 complex has been demonstrated to reduce the activity of the complex\textsuperscript{32}. ARP2 deficiency was demonstrated to influence platelet shape change, a process that is critical for normal platelet function and thus for hemostasis\textsuperscript{33}.

**Complete Actr2 deficiency is incompatible with survival**

We attempted to generate an independent Actr2 knockin (Actr2\textsuperscript{G/knockout allele (Actr2\textsuperscript{G})} by CRISPR/Cas9 technology. We were able to detect blastocysts with our prepared CRISPR materials. We have analyzed these 13 blastocysts and have found one that harbored a homozygous C to G mutation (Sup Figure 2C, first gold arrow), as well as 3 that harbored a heterozygous C to G mutation. We have also identified our engineered homozygous synonymous G to A mutation (Sup Figure 2C, second gold arrow).

The efficiency of this strategy for establishing stable CRISPR mouse lines is high, as it was reported that for every 100 embryos that underwent pronuclear co-injection, 13.3 genetically modified embryos were produced\textsuperscript{34}. Of the 300 embryos injected and implanted, we obtained 18 live mice, none of which had an Actr2 modification by either nonhomologous end joining (NHEJ) or homology directed repair (HDR). Since modifications have been shown to occur with high frequency\textsuperscript{35}, this suggests that Actr2 deficiency is incompatible with survival.

We queried the components of the ARP2/3 including ARP2, ARP3 and ARPC 1-5 for loss of function mutations. ARPC1A, ARPC2, ARPC3, ARPC4, ARPC5, ARP2 and ARP3 all had a significantly reduced or absent loss of function mutations (Table 1, ExAC, [http://exac.broadinstitute.org](http://exac.broadinstitute.org) accessed: 10/2016).
Upon analysis of Neuro2A cells, we were not able to detect complete Actr2 deficiency in any of 36 single cell colonies we successfully propagated, further suggesting that Actr2 is incompatible with survival.

VTE susceptibility genes have been refractory to a number of methods of discovery, likely due to the incomplete penetrance and variable expressivity of VTE interposed with the significant environmental risk component. This is based on the fact that although many small and large-scale genetic studies have been performed, few novel VTE susceptibility loci have been identified\(^3, 36\). Given the issue of identifying “missing heritability” in human genetics studies\(^37\), forward genetic studies using ENU mutagenesis/NGS for mutant identification can drive the discovery of novel pathways (and drug targets) contributing to disease pathology. A distinct advantage of the ENU approach over human studies is that an identified mutation will nearly always be causative, because genetically linked variants will be rare. The specific advantage of a sensitized genetic screening approach that we have adopted here means that the entire mutation space can be queried for identifying suppressors of our specific \(F5^{L/L}\ Tfpi^{+/−}\) phenotype. These types of phenotype-genotype correlations have been very difficult to form through genetic studies in humans but as described in the introduction, like our study have been successful in mice\(^38\).

The problem of endogenous strain modifier genes interfering with ENU mutation induced phenotypes has been recognized for some time, with 50% of ENU induced phenotypes being lost upon outcrossing to a different strain to introduce the genetic diversity necessary for standard gene mapping approaches\(^22, 39\). Our experimental approach affirms the use of NGS for identification of ENU induced mutants by circumventing lengthy gene mapping approaches\(^40\).
and also for enabling the identification of ENU mutants like those we have described here whose phenotype may have been obscured by “endogenous” mouse strain modifiers.

Finally and most importantly, through the design and execution of this sensitized ENU mutagenesis screen, we have identified genomic loci capable of suppressing $F5^{L/L}$ $Tfpi^{+/−}$ lethality. Specifically, our work has illustrated the importance of the TF/TFPI axis in thrombosis regulation as well as the identification of Actr2 as a prime candidate the $MfyL1$ suppressor mutant.

**Materials and methods**

**Mice**

C57BL/6J (B6, stock number 000664), 129S1/SvImJ mice (129S1, stock number 002448), were purchased from Jackson Laboratories. $F5^{L/L}$ ($F5^{tm2Dgi}/J$ stock number 004080) mice were previously generated$^{22}$. $F3$ and $Tfpi$ deficient mice were a generous gift of Dr. George Broze$^{41,42}$. $F8$ deficient mice were a generous gift of Dr. Haig Kazazian$^{43}$. All mice designated to be on the B6 background were backcrossed greater than 8 generations to B6. $F5^{L/L}$ breeding stock for genetic mapping were generated from $F5^{L}$ mice serially backcrossed greater than 12 generations to the 129S1 strain to create $F5^{L}$ congenic mice. All mice were maintained on normal chow in a specific pathogen-free (SPF) facility. All animal care and experimental procedures complied with the principles of Laboratory and Animal Care established by the National Society for Medical Research and were approved by the University of Michigan Committee on Use and Care of Animals.

**Genotyping**
DNA was isolated from tail biopsies and mice genotyped for TfpI+/- and F5L as previously described\textsuperscript{13}. Mice were genotyped for F3 deficiency using custom primers listed in Sup. Table 10. All primers were purchased from IDT, Coralville, IA.

**ENU mutagenesis and breeding**

ENU was purchased (Sigma Aldrich, St. Louis MO) in ISOPAC vials, and prepared according to the following protocol [http://pga.jax.org/enu_protocol.html](http://pga.jax.org/enu_protocol.html). A single ENU dose of 150 mg/kg was administered intraperitoneally into an initial cohort of 159 F5L/L B6 male mice (referred to as generation 0 or G0 mice). For a second cohort of 900 male F5L/L G0 mice, the protocol was changed to three weekly intraperitoneal injections of ENU (90 mg/kg). After a 10-week recovery period, each G0 mouse was bred to F5L/+ TfpI+/- mice (Figure 1B) on the B6 genetic background to produce G1 generation offspring, which were genotyped at two weeks of age. G1 mice of the F5L/L TfpI+/- genotype surviving to weaning age were considered to carry a suppressor mutation.

**Modifier gene transmission**

F5L/L TfpI+/- G1 founders were crossed to F5L/L mice on the B6 genetic background to produce G2 generation offspring. G2 mice were outcrossed to F5L/L mice on the 129S1 genetic background for 2 or more generations. Progeny testing was considered positive when the identification of F5L/L TfpI+/- offspring was greater than or equal to 2.4% of the total progeny.

**Genetic Mapping**

Genetic markers distinguishing the B6 and 129S1 strains distributed across the genome were genotyped using the Illumina GoldenGate Genotyping Universal-32 platform (Illumina, San Diego CA) at the University of Michigan DNA Sequencing Core. Linkage Analysis was performed on the Mendel platform version 14.0\textsuperscript{44} using 806 informative markers from the total
of 1449 genotyped markers. LOD scores >=3.3 were considered significant\textsuperscript{45}. The number of mice, the number of SNP markers and the LOD scores for each of the mapped pedigrees are listed in Sup. Table 5.

**Sanger sequencing of the F3 gene and analysis of candidate mutations**

Genomic DNA was extracted from mouse tail biopsies using the Gentra Puregene Tissue Kit (Qiagen, Germantown, MD). A total of 48 overlapping pairs of amplicons (primers: F3gene_1-F3gene_35; upstreamF3_1-upstreamF3_13, Sup. Table 10) were used to Sanger sequence the entire F3 gene (~11kb) and an additional ~5kb of upstream sequences on both strands. Sanger sequencing was performed at the University of Michigan Sequencing Core. For the analysis of candidate mutations, amplicons were generated harboring the nucleotide of interest using the outer primer pairs. Inner forward and reverse primers were used to bidirectionally sequence these amplicons. Sequencing chromatograms were visualized and manually scored using FinchTV.

**Estimation of F3 allelic expression**

F5\textsuperscript{L/L} Tfp1\textsuperscript{+/−} mice with one B6 allele (in cis with ENU induced variants) and one 129S1 allele at the Chr3 candidate region were outcrossed to DBA wildtype females introducing exonic B6-129S1/DBA SNPs. Five progeny from this cross (2 B6/DBA and 3 129S1/DBA allele carriers, identified by DNA genotyping) were tested for differential allelic expression. From each mouse three tissue samples (lung, liver, whole brain) were obtained as previously described\textsuperscript{22}. RNA was extracted from the tissue samples using RNeasy Plus Mini Kit (Qiagen) according to manufacturer's recommendations and reverse transcribed using SuperScript II (Invitrogen, Carlsbad, CA). cDNA corresponding to exon3-exon5 was amplified with primers F3-exon-F (5’TGCTTTCGACCACAGACAC) and F3-exon-R (5’CTGCTTCTGGGCTATTTTG), using
Gotaq Green Master Mix (Promega, Madison, WI). Primers F3-exon-F and F3-exon-R were also used to Sanger sequence the F3 exonic region.

The F3 exonic region harbors 3 known B6-129S1/DBA SNPs (rs30268372, rs30269285, rs30269288, http://www.ncbi.nlm.nih.gov/SNP/) that were used for relative expression analysis. Relative expression was estimated at SNP sites by dividing the area under the Sanger sequencing peak of one allele to another\cite{46,47}. Next, the relative expression of each SNP was compared between the B6 and 129S1 allele carrying progeny.

**Mouse whole exome sequencing**

Libraries were prepared using Agilent (Agilent Technologies, Santa Clara, CA) or NimbleGen (Roche NimbleGen, Madison, WI) mouse whole exome capture kits. 100 bp paired-end sequencing was performed on the Illumina Hiseq 2000 platform at the University of Michigan DNA Sequencing Core. A detailed overview of the whole exome sequencing pipeline is available at GitHub (https://github.com/tombergk/FVL_SUP). Briefly, sequence reads were aligned using Burrows-Wheeler Alignment software\cite{48} to the mouse reference genome (genome assembly GRCm38, Ensembl release 73). Reads were sorted and duplications removed using Picard tools (http://picard.sourceforge.net). Coverage statistics were estimated using QualiMap software\cite{49}. Variants were called across 8 samples using GATK HaplotypeCaller software\cite{50}. Standard hard filters recommended by the Broad Institute were applied using GATK VariantFiltration\cite{50} followed by an in-house developed pipeline to remove variants between the B6 and 129S1 strains, shared variants within our mouse cohort and variants in closer proximity than 200 base pairs from each other. Variants were annotated using Annovar software\cite{51} with Refseq annotation (release 61). Heterozygous variants within exonic regions with >6X coverage unique for only one mouse in the cohort were regarded as potential ENU induced variants. A
total of 125 heterozygous variants were identified as candidate suppressor mutations, using an in-house filtering pipeline\textsuperscript{46}, with 79 variants occurring within coding sequence. The number of ENU variants identified in each exome sequenced mouse varied by genealogical distance from the G1 \textit{MF5L} founder. The candidate ENU induced variants were validated by Sanger sequencing.

\textbf{Generation of Actr2 CRISPR/Cas9 targeted mice and cells}

\textit{Actr2 Targeting Sequence Design and Cloning:} CRISPR/Cas9 is widely used for genome editing. Cas9 is guided to a sequence specific cleavage site by small single-guide RNAs (sgRNA) through Watson-Crick base pairing with target DNA\textsuperscript{52-54}. We obtained expression vectors for human-codon optimized \textit{S. pyogenes} Cas9 and chimeric sgRNA (pSpCas9(BB); pX330; plasmid ID: 42230)\textsuperscript{55} as well as the puromycin resistance gene for the selection of transfected cells (pSpCas9(BB)-2A-Puro; pX459; plasmid ID:48139)\textsuperscript{56} from Addgene. Each vector was digested with \textit{BbsI} and a pair of annealed oligos (sgRNA) was independently cloned into the backbone vector as described\textsuperscript{56} and depicted in Sup. Fig 2.

The required \textit{Actr2}-specific sgRNAs were selected based on: 1) their proximity to the mutation site (< 100 base-pair (bp) away; 10 bp optimal); 2) the presence of a protospacer adjacent motif (PAM) “NGG” sequence adjacent to the sgRNA; 3) the ability to incorporate a synonymous variant within the PAM to protect the homology directed repair (HDR) donor template from Cas9-targeted degradation; and 4) the sequence must have an inverse likelihood of off-target binding score of >70 and no less than 3 mismatches within exonic regions of the genome as determined by the CRISPR design tool (\texttt{crispr.mit.edu}).

\textit{Single-Stranded DNA Donor for Homology Directed Repair:} The single-stranded DNA (ssDNA) oligo that served as an HDR donor template was ordered as an Ultramer DNA oligo
from Integrated DNA Technologies (IDT). This HDR donor template consists of a 161 bp genomic sequence homologous to a region spanning -44 to +117 bp from the splice junction of intron 6 and exon 7 of the mouse Actr2 gene (Fig. 8B). This HDR donor encodes an arginine (R) to glycine (G) mutation at the 258th amino acid position (c.772C>G) and a synonymous mutation within the PAM to prevent donor DNA cleavage by Cas9. In this design, 80 bp homology arms flank the C to G transversion mutation with the position of the double-strand break (DSB) occurring 11 – 12 bp downstream of the homology arm junction for the sgRNA.

Mouse Pronuclear Injection for Validation of sgRNA and HDR Efficiency: Microinjection of CRISPR/Cas9 constructs into fertilized eggs and then analyzing blastocyst stage embryos for toxicity and efficacy is effective for ensuring that CRISPR design and reagent preparation have been carried out effectively. Blastocysts obtained just before implantation are suitable for testing the fidelity of CRISPR/Cas9-mediated genome editing because they can be individually handled in vitro and obtained 3 days after microinjection. The University of Michigan Transgenic Animal Model Core under the direction of Dr. Thomas Saunders has completed injection of 13 pronuclei and has grown them to the blastocyst stage.

Statistical Data Analysis

Statistical differences among the potential progeny of mouse crosses were determined using the X² test. The paired t-test was used for estimating statistical differences between the weights of F5/L Tfpi+/− mice and their littermates. Relative expression differences for F3 alleles were estimated using the Wilcoxon rank-sum test. Kaplan Meier analysis was used to assess significance for putative suppressors identified by exome sequencing. Probability of survival was calculated and plotted using Medcalc.
Acknowledgements

This research was supported by NIH grant P01-HL057346 (to D. Ginsburg). R.J. Westrick was generously supported by the following American Heart Association grants: Predoctoral Fellowship 0315212Z, Innovative Research Grant 10IRG3740035, and Scientist Development Grant 12SDG12140000. K. Tomberg is an International Fulbright Science and Technology Fellow and the recipient of an American Heart Association Predoctoral Fellowship Grant. A.E. Siebert is the recipient of an Oakland University Provost Award. M.A. Brake is the recipient of an American Heart Association Undergraduate Fellowship 15UFEL25700204. D. Ginsburg in an Investigator of the Howard Hughes Medical Institute. We gratefully acknowledge expertise of the Transgenic Animal Model Core of the University of Michigan’s Biomedical Research Core Facilities for assistance with generating the Actr2 CRISPR/Cas9 mice used in this study.
References


9. Recommendations from the egapp working group: Routine testing for factor v leiden (r506q) and prothrombin (20210g>a) mutations in adults with a history of idiopathic venous thromboembolism and their adult family members. Genet Med. 2011;13:67-76


Figure legends

Figure 1: *F8* deficient thrombosuppression and design of the *Leiden* ENU mutagenesis screen. **A.** The mating scheme and observed distributions of the *F5*<sup>L+/+</sup> *Tfpi<sup>+/−</sup> F8 deficiency rescue experiments. *F8* <i>X</i> results in incompletely penetrant suppression of the *F5*<sup>L+/+</sup> *Tfpi<sup>+/−</sup>* phenotype. **B.** The mating scheme and observed distribution of the *Leiden* screen. *F5*<sup>L+/+</sup> *Tfpi<sup>+/−</sup>* male mice were mutagenized with either 1 x 150mg/kg or 3 x 90 mg/kg ENU and bred with non-mutagenized *F5*<sup>L+/−</sup> females. Sixteen and 83 *F5*<sup>L+/−</sup> *Tfpi<sup>+/−</sup>* progeny, respectively were observed in each of the dosing regimens, with over twice the rate of *F5*<sup>L+/−</sup> *Tfpi<sup>+/−</sup>* survivors in the progeny of the 3 x 90 mg/kg treated mice. **C.** Though a trend toward worse survival among female *F5*<sup>L+/−</sup> *Tfpi<sup>+/−</sup>* putative suppressor mice, this difference was not significant (p=0.114). **D and E.** *F5*<sup>L+/−</sup> *Tfpi<sup>+/−</sup>* putative suppressor mice were significantly smaller than their non-*F5*<sup>L+/−</sup> *Tfpi<sup>+/−</sup>* littermates in both the pure B6 and mixed B6-129S1 genetic backgrounds.

Figure 2: The MF5L6 suppressor locus maps to Chr 3. **A.** Linkage analysis for the MF5L6 line. The Chr 2 locus (LOD score=9.81) includes the *Tfpi* gene. The Chr 3 peak had the highest LOD score in the Chr3 subregion:117.3-124.8Mb (maximum LOD=4.49, 1 LOD interval). The most significant linkage signal (Chr 2, LOD score=9.81) spans the *Tfpi* locus and was generated by markers tightly linked to the *Tfpi* allele. Due to the backcross history of our *Tfpi* mouse colony, these mice have a complicated *Tfpi* congenic interval containing 129S1 and B6 alleles linked to the knockout allele. Since these alleles are “passenger alleles”, they will always be heterozygous as a result of our mating strategy (van vlijmen ref). **B.** The mating scheme and observed distribution of offspring to test *F3* deficiency as a suppressor of *F5*<sup>L+/−</sup> *Tfpi<sup>+/−</sup>. *F3*<sup>+/−</sup> results in incompletely penetrant suppression of the *F5*<sup>L+/−</sup> *Tfpi<sup>+/−</sup>* phenotype. **C.** The Chr 3 candidate interval (chr3:117.3-124.8 Mb) contains 38 refseq annotated genes, including *F3*.

Figure 3: Discovery and validation of *Actr2* R258G as a thrombosis suppressor gene by NGS. **A.** Kaplan-Meier survival plot for *F5*<sup>L+/−</sup> *Tfpi<sup>+/−</sup>* mice with and without the *Actr2* mutation. *F5*<sup>L+/−</sup> *Tfpi<sup>+/−</sup> *Actr2<sup>+/G</sup>* exhibit significantly better survival than *F5*<sup>L+/−</sup> *Tfpi<sup>+/−</sup> *Actr2<sup>+/+</sup> (n = 35 mice). **B.** ARP2 amino acid R258 is highly conserved in animals, plants and fungi.
A

\[ p = 1.7 \times 10^{-6} \]

B

\textit{Actr2} gene

\begin{align*}
1 & \quad 7 & \quad 9 \\
PQFLAEPAEFREGGKIIRGE & \quad M.\textit{musculus} & \quad \text{ENU mutation} \\
PQFLAEPAEFREGGKIIRGE & \quad R.\textit{norvegicus} \\
PQFLAEPAEFREGGKIIRGE & \quad O.\textit{cuniculus} \\
PQFLAEPAEFREGGKIIRGE & \quad H.\textit{sapiens} \\
PQFLAEPAEFREGGKIIRGE & \quad P.\textit{troglodytes} \\
PQFLAEPAEFREGGKIIRGE & \quad G.\textit{gorilla} \\
PQFLAEPAEFREGGKIIRGE & \quad S.\textit{scrofa} \\
PQFLAEPAEFREGGKIIRGE & \quad B.\textit{taurus} \\
PQFLAEPAEFREGGKIIRGE & \quad A.\textit{melanoleuca} \\
PQFLAEPAEFREGGKIIRGE & \quad E.\textit{caballus} \\
PQFLAEPAEFREGGKIIRGE & \quad L.\textit{africana} \\
PQFLAEPAEFREGGKIIRGE & \quad M.\textit{Lucifugus} \\
PQFLAEPAEFREGGKIIRGE & \quad D.\textit{novemcinctus} \\
PQFLAEPAEFREGGKIIRGE & \quad O.\textit{anatinus} \\
PQFLAEPAEFREGGKIIRGE & \quad G.\textit{gallus} \\
PQFLAEPAEFREGGKIIRGE & \quad A.\textit{carolinensis} \\
PQFLAEPAEFREGGKIVRGD & \quad X.\textit{tropicalis} \\
PQFLAEPAFREGGKIVRGD & \quad T.\textit{rubripes} \\
PQFLAEPAFREGGKIVRGD & \quad G.\textit{aculeatus} \\
PQFLAEPAFREGGKIVRGD & \quad D.\textit{rerio} \\
PQFLAEPAFREGGKIVRGD & \quad D.\textit{melanogaster} \\
PQFLAEPAFREGGLIVRGD & \quad C.\textit{elegans} \\
PQFLCEPAFREGGKIVRGD & \quad S.\textit{cerevisiae} \\
PQFLAEPAQFREGTVKIVRGD & \quad A.\textit{thaliana}
\end{align*}