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

Randal J. Westrick[a,b,c,i], Kärt Tomberg[c,e,i], Amy E. Siebert[a,i], Guojing Zhu[c], Mary E. Winn[d], Sarah L. Dobies[c], Sara L. Manning[c], Marisa A. Brake[a], Audrey C. Cleuren[c], Linzi M. Hobbs[a], Lena M. Mishack[a], Alexander Johnston[a], Emilee Kotnik[c], David R. Siemieniak[f], Jishu Xu[e], Jun Z. Li[e], Thomas L. Saunders[g] and David Ginsburg[c,e,f,h]

[a]Oakland University Department of Biological Sciences

[b]Oakland University Center for Data Science and Big Data Analysis

[c]Life Sciences Institute, University of Michigan

[d]Bioinformatics and Biostatistics Core, Van Andel Research Institute

[e]Department of Human Genetics, University of Michigan

[f]Howard Hughes Medical Institute, University of Michigan

[g]Transgenic Animal Model Core, University of Michigan

[h]Departments of Internal Medicine and Pediatrics, University of Michigan

[i]These authors contributed equally to this work

Corresponding author:

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.

[48 pages, 227 words in abstract, 5,340 words, 30,327 characters not including abstract, title page, figures and references]

Abstract

Factor V Leiden ($F5^{L}$) is a common genetic risk factor for venous thromboembolism in humans. We conducted a sensitized ENU mutagenesis screen for dominant thrombosuppressor genes based on perinatal lethal thrombosis in mice homozygous for $F5^{L}$ ($F5^{L/L}$) and haploinsufficient for tissue factor pathway inhibitor ($Tfpi^{+/-}$). *F8* deficiency enhanced survival of $F5^{L/L}$ $Tfpi^{+/-}$ mice, demonstrating that $F5^{L/L}$ $Tfpi^{+/-}$ lethality is genetically suppressible. ENU-mutagenized $F5^{L/L}$ males and $F5^{L/+}$ $Tfpi^{+/-}$ females were crossed to generate 6,729 progeny, with 98 $F5^{L/L}$ $Tfpi^{+/-}$ offspring surviving until weaning. Sixteen lines exhibited transmission of a putative thrombosuppressor to subsequent generations, with these lines referred to as *MF5L* (**M**odifier of **F**actor **5** **L**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, haploinsufficiency for *F3* ($F3^{+/-}$) suppressed $F5^{L/L}$ $Tfpi^{+/-}$ lethality. Whole exome sequencing in *MF5L12* identified an *Actr2* gene point mutation (p.R258G) as the sole candidate. Inheritance of this variant is associated with suppression of $F5^{L/L}$ $Tfpi^{+/-}$ lethality ($p=1.7x10^{-6}$), suggesting that $Actr2^{p.R258G}$ is thrombosuppressive. CRISPR/Cas9 experiments to generate an independent *Actr2* knockin/knockout demonstrated that *Actr2* haploinsufficiency is lethal, supporting a hypomorphic or gain of function mechanism of action for $Actr2^{p.R258G}$. Our findings identify *F8* and the *Tfpi/F3* axis as key regulators in determining thrombosis balance in the setting of $F5^{L}$ and also suggest a novel role for *Actr2* in this process.

Significance Statement (120 words max):

Venous thromboembolism (VTE) is a common disease characterized by the formation of abnormal blood clots. Inheritance of specific genetic variants, such as the Factor V Leiden polymorphism, increases VTE susceptibility. However, only ~10% of people inheriting Factor V Leiden will develop VTE, suggesting the involvement of other genes that are currently unknown. By inducing random genetic mutations into mice with a genetic predisposition to VTE, we identified two genomic regions that reduce VTE susceptibility. The first includes the tissue factor gene and its role was confirmed by analyzing mice with an independent tissue factor gene mutation. The second contains a mutation in the *Actr2* gene. These findings identify critical genes for the regulation of blood clotting risk.

Keywords:

Venous thromboembolism, Factor V Leiden, ENU mutagenesis

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*, *F2 F5*, *F11*, *FGG* and *PROCR* as thrombosis susceptibility genes, with only two additional novel loci, TSPAN15 and SLC44A2 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 average allele frequency of 3.5% in the European population(6-8). $F5^L$ is estimated to account for up to 25% of the genetically-attributable thrombosis risk in these populations(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, 9), limiting the clinical utility of $F5^L$ genotyping in the management of VTE(10).

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(9, 11, 12). 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, 9).

Mice carrying the orthologous $F5^L$ mutation exhibit a mild to moderate prothrombotic phenotype closely mimicking the human disorder(13). We previously reported a synthetic lethal

interaction between $F5^{L}$ homozygosity ($F5^{L/L}$) and hemizygosity for tissue factor pathway inhibitor ($Tfpi^{+/-}$)(14). Nearly all mice with this lethal genotype combination ($F5^{L/L}$ $Tfpi^{+/-}$) succumb to widespread, systemic thrombosis in the immediate perinatal period(14).

ENU mutagenesis in mice has been used effectively to identify novel genes involved in a number of biological processes(15, 16). 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(17). 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(18). Other successful sensitized screens include analysis of mouse mutants predisposed to diabetic nephropathy(19), a screen in *Sox10* haploinsufficent mice identifying the *Gli3* gene as a modifier of neurochristopathy(20) and identification of a mutation in the *c-Myb* gene as a dominant modifier for platelet count in *Mpl* deficient mice (congenital thrombocytopenia)(21). We now report the results of a dominant, sensitized ENU mutagenesis screen for suppressors of $F5^{L/L}$ $Tfpi^{+/-}$ dependent lethal thrombosis.

Results and Discussion

F8 deficiency suppresses $F5^{L/L}$ $Tfpi^{+/-}$ lethality

X-linked hemophilia A results in a moderate to severe bleeding disorder in humans and is caused by mutations in the *F8* gene. To test whether the $F5^{L/L}$ $Tfpi^{+/-}$ lethal thrombotic phenotype is suppressible by hemophilia A in mice, triple heterozygous $F5^{L/+}$ $Tfpi^{+/-}$ $F8^{+/-}$ female mice were generated and crossed to $F5^{L/L}$ male mice (Fig. 1A). One quarter of conceptuses are expected to carry the $F5^{L/L}$ $Tfpi^{+/-}$ genotype, with half of the total expected male conceptuses completely *F8*

deficient (*F8*$^{-}$). Thus, 1/16th of the overall offspring from this mating are expected to be *F5*$^{L/L}$ *Tfpi*$^{+/-}$ *F8*$^{-}$ (males). Similarly, 1/16th of the progeny should be *F5*$^{L/L}$ *Tfpi*$^{+/-}$ *F8*$^{+/-}$ (females). A total of 163 progeny from this cross were genotyped at weaning, resulting in 8 *F5*$^{L/L}$ *Tfpi*$^{+/-}$ *F8*$^{-}$ male mice observed (and 0 *F5*$^{L/L}$ *Tfpi*$^{+/-}$ *F8*$^{+}$, p=0.02) and 2 *F5*$^{L/L}$ *Tfpi*$^{+/-}$ *F8*$^{+/-}$ female mice (and 1 *F5*$^{L/L}$ *Tfpi*$^{+/-}$ *F8*$^{+/+}$, p=0.9). These results demonstrate that *F5*$^{L/L}$ *Tfpi*$^{+/-}$ thrombosis is genetically suppressible by *F8* deficiency with nearly complete penetrance in *F8*$^{-}$ male mice and are consistent with human studies demonstrating *F8* level as an important VTE risk factor(22).

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 mutant phenotypes(23), 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(23). This is consistent with the ~20-30 functionally significant mutations per G1 mouse expected with this ENU mutagenesis protocol(24, 25). One quarter of G1 embryos from this cross are expected to carry the synthetic lethal *F5*$^{L/L}$ *Tfpi*$^{+/-}$ genotype. Out of a total of 6,729 G1 mice screened at weaning, the 98 live *F5*$^{L/L}$ *Tfpi*$^{+/-}$ mice (45 females, 53 males) represented 4.4% of the 2,210 embryos expected with this genotype. Survival data were collected for 57 of the *F5*$^{L/L}$ *Tfpi*$^{+/-}$ G1 mice, with 34 living past 70 days of age (precise dates of death were not available for the remaining 41 mice). No significant sex-specific differences in survival were observed (Fig. 1C).

Heritability for each of the 44 G1 putative suppressor mutants who lived to breeding age 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. Out of the original 98 surviving $F5^{L/L}$ $Tfpi^{+/-}$ G1 mice, 75 produced no offspring surviving to weaning, either due to infertility or the above mentioned early lethality, with >50% of these mice (37 of 75) exhibiting a grossly runted appearance. Approximately half of the $F5^{L/L}$ $Tfpi^{+/-}$ G1 mice that attained breeding age (23/44) produced 1 or more G2 progeny surviving to weaning, 7 (2 males and 5 females) produced no $F5^{L/L}$ $Tfpi^{+/-}$ G2s, including 4 G1s with 8 or more offspring of other genotypes. Sixteen $F5^{L/L}$ $Tfpi^{+/-}$ G1 mice produced one or more $F5^{L/L}$ $Tfpi^{+/-}$ progeny when bred to B6 $F5^{L/L}$ mice (see Methods). These 16 potential thrombosuppressor mouse lines are designated *MF5L1-16*. The number of total progeny, genotypic distribution, and penetrance of the $F5^{L/L}$ $Tfpi^{+/-}$ mice in each line are listed in Table S1. Within these suppressor lines, mice with the $F5^{L/L}$ $Tfpi^{+/-}$ genotype were ~30% smaller than their $F5^{L/L}$ littermates at the time of weaning ($p<2.2x10^{-16}$, Fig. 1D), with this difference maintained after outcrossing to the 129S1 strain (Fig. 1E).

Previous reports based on gene function in the specific locus test estimate an ENU-induced mutation rate of 1/700 loss of function mutations per locus for the ENU dosing regimen used here(26). This mutation rate predicts that our screen of 6,729 G1 progeny (2,210 $F5^{L/L}$ $Tfpi^{+/-}$ expected) should have produced ~3 mutations per gene averaged over the entire genome, with 54% of these mutations expected to be null(27), resulting in 1.5X genome coverage for loss of function mutations.

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

In order to map putative ENU-induced suppressor mutations, surviving $F5^{L/L}$ $Tfpi^{+/-}$ mice were intercrossed with $F5^{L/L}$ mice that had been backcrossed onto the 129S1/SvIMJ strain

(129S1). Crosses between $F5^{L/L}$ and $F5^{L/+}$ $Tfpi^{+/-}$ mice (both $F5^{L}$ and $Tfpi^{-}$ backcrossed > 12 generations onto 129S1) confirmed the lethality of the $F5^{L/L}$ $Tfpi^{+/-}$ genotype on the 129S1 background (Table S2). The 4 lines containing the largest number of genetically informative B6-129S1 mixed background $F5^{L/L}$ $Tfpi^{+/-}$ offspring (*MF5L1, 6, 9* and *16*) were used for gene mapping. Though the *MF5L1*, *MF5L9* and *MF5L16* lines 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 ≳3 (maximum LODs for *MF5L1=1.15*, *MF5L9=2.5* and *MF5L16=1.61*). 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(9, 28) and confound linkage analysis(29). Consistent with this hypothesis, we have previously documented poorer survival to weaning in mixed B6-129S1 $F5^{L/L}$ mice compared to littermates(13). We extended these observations by the analysis of additional $F5^{L/+}$ and $F5^{L/L}$ littermates, with mice of the $F5^{L/L}$ genotype demonstrating a 50% reduction in survival in the 129S1 versus B6 strain backgrounds (Table S2).

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 on the mixed B6-129S1 background and therefore useful for linkage analysis, See Table S1). Genome-wide SNP genotyping was performed on DNA from these 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 Fig. 2A Legend and Methods). A single locus with a significant LOD score of 4.49 was identified on chromosome (Chr) 3, with the 1 LOD interval (117.3-124.8Mb) containing 43 refseq annotated genes (Fig. 2B).

The *F3* gene located within this interval (Chr3:121.7 Mb) (Fig. 2B) 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. To test *F3* as a candidate suppressor of the $F5^{L/L}$ $Tfpi^{+/-}$ phenotype, an independent *F3* null allele was introduced and triple heterozygous $F5^{L/+}$ $Tfpi^{+/-}$ $F3^{+/-}$ mice were crossed to $F5^{L/L}$ B6 mice (Fig. 2C). Of 273 progeny genotyped at weaning, 13 $F5^{L/L}$ $Tfpi^{+/-}$ $F3^{+/-}$ were observed (and 1 $F5^{L/L}$ $Tfpi^{+/-}$ $F3^{+/+}$, $p=9.7\times10^{-5}$). 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, although with incomplete penetrance (33%) that also differs by sex (10% for males and 67% for females). In contrast, the *MF5L6* line exhibited an overall penetrance of 72.4%, with similar male/female penetrance. Gender-specific differences in venous thrombosis rates have previously been reported, including contributions from oral contraceptives and hormone replacement therapy(30-32). This difference in penetrance could be due to 129S1 strain effects in the *MF5L6* line or differences between a *F3* regulatory mutation in *MF5L6* compared to the *F3* loss of function allele used here.

Whole exome sequencing data analysis of a $F5^{L/L}$ $Tfpi^{+/-}$ mouse from *MF5L6* failed to identify an ENU variant in *F3* or in any other genes in the nonrecombinant interval, or more broadly on the entire Chr3. Although additional ENU variants were identified on other chromosomes, none co-segregated with the survival phenotype in line *MF5L6* (Table S3). Sanger sequencing analysis of the full set of *F3* exons and introns, as well as 5kb upstream of exon 1, also failed to identify an ENU-induced mutation. In addition, 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 (Fig. S1).

Taken together, these data suggest that an ENU-induced *F3* regulatory mutation outside of the sequenced segment may be responsible for thrombosuppression in *MF5L6*, although we cannot exclude a regulatory mutation in another gene. Nonetheless, our findings demonstrate that *F3*/*Tfpi* balance plays a key role in thrombosis in the mouse, particularly in the setting of $F5^{L}$, 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 an index $F5^{L/L}$ $Tfpi^{+/-}$ mouse (from the G2-G5 generation) 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 (Table S4). A total of 125 heterozygous variants were identified as candidate suppressor mutations, with 79 variants affecting protein sequence (Table S3). Of the total mutations, 54.4% 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 (35.3%), while C/G→G/C transversions were the least represented (2.5%). This spectrum of mutations is consistent with previously published ENU experiments(33). Variants exhibiting no recombination with the *Tfpi* locus on Chr2 (17 variants) were excluded from further analysis (See Methods). Sanger sequencing confirmation was performed for 62 variants, including all nonsynonymous and stopgain mutations. These variants were then checked for parent of origin (either the G1 mutagenized progeny or its non-mutagenized mate) as well as the original mutagenized G0 male.

Forty-seven of these variants were identified in the G1 mouse but not in the G0 or non-mutagenized parent, consistent with ENU-induced mutations. The remaining 15 mutations were either NGS sequencing errors (11/15), de novo (2/15), or transmitted from the non-mutagenized parent (2/15) (Table S3).

Each SNV was analyzed in additional *MF5L* mice from the line in which it was identified. None of the thrombosuppressive exonic ENU-induced variants identified in lines *MF5L1, 5, 6, 8, 9, 11* and *16* segregated with the lethal phenotype as tested by Kaplan-Meier analysis using a significance threshold of $p<0.05$(34). Of the 7 candidate ENU-induced SNVs identified from whole exome sequence analysis for the *MF5L12* line, 1 was an NGS sequencing error and 6 were validated by Sanger sequencing as consistent with ENU-induced mutations in the G0 mice (Table S3). For each of these 6 SNVs, co-segregation with the survival phenotype was tested by Kaplan-Meier analysis of the first 31 $F5^{L/L}$ $Tfpi^{+/-}$ mice generated from the *MF5L12* line. Only one variant, a nonsynonymous SNV in the *Actr2* gene (c.772C>G, p.R258G, $Actr2^{G}$), demonstrated a significant survival advantage when co-inherited with the $F5^{L/L}$ $Tfpi^{+/-}$ genotype ($p=1.7x10^{-6}$) (Fig. 3A).

***Actr2* as a candidate thrombosuppressor gene**

The *Actr2* gene encodes the ARP2 protein, which is an essential component of the Arp2/3 complex(35). ARP2, along with ARP3 and five other independent protein subunits (ARPC1-5), forms the evolutionarily conserved seven-subunit Arp2/3 complex(36). Arp2/3 is a major component of the actin cytoskeleton and is found in most eukaryotic cells including platelets(37). Arp2/3 binds to the sides of actin filaments and initiates the growth of a new filament, leading to the creation of branched actin networks that are important for many cellular processes(38). Loss of Arp2/3 function can have severe consequences, as illustrated by the embryonic lethality of

mice homozygous for an ARP3 hypomorph(39). In hemostasis, the Arp2/3 complex is necessary for actin-dependent platelet cytoskeletal remodelling events, which are essential for platelet activation and degranulation(40-42). 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), as well as in plants and fungi (Fig. 3B). In addition, no variants at this position have been identified to date in over 120,000 human alleles(43).

Actr2 hemizygosity is incompatible with survival

We attempted to generate an independent *Actr2* knockin (*Actr2*^{G}) allele by CRISPR/Cas9 genome editing. CRISPR/Cas9-induced homology directed repair (HDR)-mediated DNA integration within the *Actr2* gene was tested in the Neuro-2a (N2a) cell line using two single guide RNAs (sgRNAs) and their respective single-stranded DNA (ssDNA) donor sequences for HDR-mediated knockin (Fig. S2A and B, Table S5). In post-selected heterogenous N2a cells, we observed integration of the variant containing DNA sequence when using sgRNA #1 (Fig. S2C).

We next injected 30 blastocysts and analyzed the 13 surviving to the 60-cell stage. Genotyping of the 13 blastocysts produced from the sgRNA #1 injected oocytes exhibited a high degree of mosaicism, with all 13 positive for DNA cleavage events and 5 (38%) positive for *Actr2*^{G} substitution (Fig. S3). However, injection of 305 fertilized eggs and subsequent transfer of 275 embryos into 9 foster mothers resulted in the generation of only 18 surviving pups at weaning, all of which were wildtype for the *Actr2* allele. Taken together with the high *Actr2*-targeting efficiency (see Fig. S3 and Methods) these data suggest that heterozygous loss of function for *Actr2* may be incompatible with survival to term.

Consistent with this hypothesis, human sequencing data from the Exome Aggregation Consortium (ExAC), which includes 60,706 individual exomes, reports a loss of function intolerance for *ACTR2* of 0.997(43). *ACTR2* mutations have not been previously associated with human disease (https://omim.org/entry/604221)(44), again consistent with early embryonic lethality. In addition, out of 373,692 mouse ENU-induced mutations listed in the Mutagenetix website, only 16 are located in the *Actr2* gene, with no predicted loss of function mutations (https://mutagenetix.utsouthwestern.edu/)(45). Taken together, these data strongly suggest that haploinsufficiency for *Actr2* is not tolerated in humans or mice. The viability of $Actr2^{+/G}$ mice suggests that the $Actr2^{G}$ allele is either hypomorphic or a unique gain of function mutation distinct from simple haploinsufficiency. Similarly, analysis of ExAC data suggests that 4 of the 6 other members of the Arp2/3 complex are intolerant to heterozygous loss of function in humans(43). Thus, the high efficiency of CRISPR/Cas9 in generating compound heterozygous loss of function variants together with the less efficient $Actr2^{G}$ substitution likely explains the outcome of our *Actr2* genome editing experiments.

The identification of novel factors involved in the regulation of hemostasis is challenging, as genes leading to marked shifts in hemostatic balance resulting in either severe bleeding or thrombosis are straightforward to identify clinically in humans, whereas subtle shifts are likely to escape detection given the multiple layers of buffering built into the complex hemostatic system(46). The dominant sensitized suppressor screen reported here was undertaken to identify genes for which modest ($\leq$50%) reduction in function would significantly shift overall hemostatic balance. Such loci represent likely candidates for common human variation contributing to thrombosis and bleeding disorders. Gene variants with subtle yet significant antithrombotic effects represent attractive therapeutic targets because of a potentially wide

therapeutic window with few unintended side effects. In contrast, a recessive screen would have been expected to identify known Mendelian bleeding disorders such as hemophilia A (*F8* deficiency), as confirmed by the data in Figure 1A. The finding of 98 $F5^{L/L}$ $Tfpi^{+/-}$ mice carrying putative thrombosis suppressor mutations (at an estimated 1.5X genome coverage) suggests that subtle alterations at a number of loci are capable of suppressing the $F5^{L/L}$ $Tfpi^{+/-}$ lethal thrombotic phenotype. The complex strain-specific genetic modifiers that confounded the genetic linkage analysis are consistent with this model. Nonetheless, our findings illustrate the particular importance of the *F3*/*Tfpi* axis in thrombosis regulation (especially in the setting of $F5^{L}$), as well as the identification of *Actr2* and the Arp2/3 complex as another potentially sensitive regulatory pathway for maintaining hemostatic balance.

Materials and methods

Mice

C57BL/6J (B6, stock number 000664), DBA2/J (DBA, stock number 000671), 129S1/SvImJ mice (129S1, stock number 002448), and B6D2F1 mice (stock number 100006) were purchased from the Jackson Laboratory. $F5^{L/L}$ ($F5^{tm2Dgi}$/J stock number 004080) mice were previously generated(22). *F3* and *Tfpi* deficient mice were a generous gift of Dr. George Broze(47, 48). *F8* deficient mice were a generous gift of Dr. Haig Kazazian(49). All mice designated to be on the B6 background were backcrossed greater than 8 generations to B6. $F5^{L/L}$ breeding stock for the 129S1 modifier gene crosses and genetic mapping experiments were generated from $F5^{L}$ mice serially backcrossed greater than 12 generations to the 129S1 strain to create B6-129S1 $F5^{L}$ congenic mice. $Tfpi^{-}$ breeding stock for the 129S1 modifier gene crosses were generated by serially backcrossed greater than 12 generations to the 129S1 strain to create B6-129S1 $Tfpi^{+/-}$ congenic mice. G1 suppressor mutant heritability was evaluated by a progeny

test backcross to B6 $F5^{L/L}$ mice. Among the 16 mice able to produce surviving $F5^{L/L}$ $Tfpi^{+/-}$ offspring, males were overrepresented (4 female, 12 male), likely because of larger numbers of offspring resulting from breeding to multiple female partners. These 16 potential thrombosuppressor mouse lines were crossed onto 129S1 to generate suppressor lines of genetically informative progeny for genetic mapping. The University of Michigan Institutional Committee on the Use and Care of Animals (IACUC) approved all experiments using mice (protocol numbers PRO00007371, PRO00005191 and PRO00005913).

Genotyping

DNA was isolated from tail biopsies and mice were genotyped for $Tfpi^{+/-}$ and $F5^{L}$ as previously described(14). Mice were genotyped for *F3* deficiency using custom primers listed in Table S6. All primers were purchased from Integrated DNA Technologies (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. A single ENU dose of 150 mg/kg was administered intraperitoneally into an initial cohort of 159 $F5^{L/L}$ B6 male mice (referred to as generation 0 or G0 mice). For a second cohort of 900 male $F5^{L/L}$ 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 $F5^{L/+}$ $Tfpi^{+/-}$ mice (Fig. 1B) on the B6 genetic background to produce G1 generation offspring, which were genotyped at two weeks of age. G1 mice of the $F5^{L/L}$ $Tfpi^{+/-}$ genotype surviving to weaning age (three weeks of age) were considered to carry a suppressor mutation.

Modifier gene transmission

F5$^{L/L}$ *Tfpi*$^{+/-}$ G1 founders were crossed to *F5*$^{L/L}$ mice on the B6 genetic background to produce G2 generation offspring. G2 mice were outcrossed to *F5*$^{L/L}$ mice on the 129S1 genetic background for 2 or more generations.

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(50) using 806 informative markers from the total of 1449 genotyped markers. LOD scores >=3.3 were considered significant(51). Chrs 1 and 2 contained the *F5* and *Tfpi* genes, respectively, and therefore these chromosomes were excluded from further analysis. The number of mice and the LOD scores for each of the mapped pedigrees are listed in Table S1.

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, Table S6) 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 a custom outer primer pair. Inner forward and reverse primers were used to bidirectionally sequence these amplicons. Sequencing chromatograms were visualized and manually scored using FinchTV (PerkinElmer, Waltham, MA).

Estimation of *F3* allelic expression

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. $F5^{L/L}$ $Tfpi^{+/-}$ 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. Three tissue samples (lung, liver, whole brain) were obtained from each mouse as previously described(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 exon 3-exon 5 was amplified with primers F3-exon-F and F3-exon-R 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.

Relative expression was estimated at SNP sites by dividing the area under the Sanger sequencing peak of one allele to another (52, 53). 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 base pair (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(54) 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(55). Variants were called across 8 samples using GATK HaplotypeCaller software(56). Standard hard filters recommended by the Broad Institute were applied using GATK VariantFiltration(56) 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 bp from each other. Variants were annotated using Annovar software(57) 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(52), 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 *MF5L* founder. The candidate ENU induced variants were confirmed by Sanger sequencing.

Generation of *Actr2* CRISPR/Cas9 targeted mice and cells

Actr2 Targeting Sequence Design and Cloning: The pX459 (pSpCas9(BB)-2A-Puro; plasmid ID:48139)(58) biscistronic expression vector for human-codon optimized *S. pyogenes* Cas9, chimeric sgRNA and puromycin resistance gene was obtained from Addgene. The vector was digested with *BbsI* and a pair of annealed oligos (sgRNA) was independently cloned into the backbone vector as described(58) and depicted in Figure S2A. Colonies were screened for successful guide insertion by *BbsI*/*AgeI* double digestion to discriminate between positive and negative clones. Clones that displayed only a ~8.5 kb linearized *AgeI* digested fragment were used in downstream applications.

The *Actr2*-specific sgRNA sequences used in these experiments are listed in Table S5. sgRNAs were selected based on: 1) their proximity to the mutation site (< 100 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 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 (crispr.mit.edu).

Single-Stranded DNA Donor for Homology Directed Repair: The ssDNA oligonucleotides that served as HDR donor templates were ordered as Ultramer DNA oligos from IDT and are listed in Table S5. The 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. S2B). The 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 or 17 – 18 bp upstream of the homology arm junction for sgRNA #1 and #2 respectively.

***Transfection of Neuro-2a cells for Validation of sgRNA and HDR Efficiency*:** Mouse N2a cells (ATCC® CCL-131TM) were routinely cultured in EMEM (Hyclone, Logan, UT), supplemented with 10% (v/v) fetal bovine serum (FBS) (Hyclone) and incubated at 37˚C in the presence of 5% CO2. N2a cells (passage 4) were seeded in triplicate into 24-well plates at a density of 1 x 10^5 cells per well in 0.5 ml EMEM containing 10% FBS 24 h prior to transfection. Cells in each well were co-transfected for 24 h with 0.5 µg pX459 plasmid (expressing either sgRNA #1 or #2) and 0.5 µg HDR donor template using 0.75 µl lipofectamine 3000 (Invitrogen/Thermo Fisher

Scientific, Waltham, MA) diluted in Opti-Mem I media (Gibco/Thermo Fisher Scientific) following manufacturer's instruction. Transfected cells were isolated by antibiotic selection using 2 μg/ml Puromycin Dihydrochloride (Gibco/Thermo Fisher Scientific) for 72 h. Transfected cells were passaged once prior to harvesting for genomic DNA extraction using the Purelink Genomic DNA mini kit (Invitrogen/Thermo Fisher Scientific) according to manufacturer's protocol.

Mouse Pronuclear Injection: Pronuclear microinjection was carried out essentially as described in Brinster et al. (59). Following pronuclear microinjection mouse zygotes were cultured *in vitro* to the blastocyst stage prior to DNA extraction and analysis for the presence of indels and mutations. Establishing stable CRISPR mouse lines by pronuclear co-injection of Cas9 mRNA, sgRNA and oligo donors is highly efficient, as it was reported that for every 100 embryos that underwent this process, ~13 genetically modified embryos were produced(61). CRISPR/Cas9 *Actr2*-edited embryos and mice were generated in collaboration with the University of Michigan Transgenic Animal Model Core (TAMC). A premixed solution of 5 ng/μl of pX459 plasmid containing sgRNA #1 targeting *Actr2* exon 7 and 10 ng/μl HDR donor template was prepared in RNAse-free microinjection buffer (10 mM Tris-HCl [pH 7.4], 0.25 mM EDTA) and microinjected into the male pronucleus of fertilized mouse eggs obtained from the mating of B6 male mice with superovulated B6 female mice. Microinjected eggs were transferred to pseudopregnant B6DF1 female mice.

Genome Extraction from Blastocyst Embryos: Mouse blastocyst DNA extraction was performed based on the method described by Sakurai *et al*.(60) Briefly, 60-cell expanded blastocysts cultured for 3 days *in vitro* at the TAMC were individually collected into 0.2 ml tubes in 10 μl of ultrapure water. Ten microliters of 2X blastocyst extraction buffer (100 mM

Tris-HCl [pH 8.3], 100mM KCl, 0.02% gelatin, 0.45% Tween-20 supplemented with 60 µg/ml yeast tRNA and 125 µg/ml Proteinase K) was added and the samples were incubated at 56°C for 10 min followed by 95°C for 10 min and immediately placed on ice to prevent heteroduplex formation. Crude lysates were stored at -20°C.

Multiplex PCR Genotyping: We designed a two-reaction multiplex PCR strategy sensitive enough to detect wildtype (WT) and R258G alleles that differ by a single nucleotide. Common ACTR2_OF forward and ACTR2_OR reverse primers were used in separate reactions with the WT-specific reverse primer (ACTR2_WT-R) or the R258G-specific reverse primer (ACTR2_MUT-R) that only differ in the -1 position on the 3'-end (Table S6). Inclusion of the common primers (659 bp product) provided amplification competition and acted as a positive PCR reaction control. Amplification of the 318 bp product using ACTR2_WT-R or ACTR2_MUT-R allele-specific primers indicates the presence of the wildtype or mutant allele respectively (Figs. S2D and S3A).

***SURVEYOR Nuclease Assay*:** The genomic region flanking the CRISPR target site for *Actr2* was PCR amplified using genomic primers ACTR2_OF and ACTR2_OR (Table S6). Unpurified PCR products (30 µl) were subjected to a denaturing and reannealing process to enable heteroduplex formation: 95°C for 10 min; 95°C to 85°C ramping at -2°C/s; 85°C to 25°C at -0.3°C/s; 25°C for 1 min and 4°C hold. After reannealing, ~250 ng DNA products were treated with SURVEYOR nuclease and SURVEYOR enhancer S (IDT) following the manufacturer's recommended protocol for GoTaq DNA polymerase amplified products. Digested and undigested (Cut/Uncut) products were analyzed by standard gel electrophoresis using 2.0% TAE agarose gels containing ethidium bromide and were imaged with a Chemi-Doc Touch gel imaging system (Bio-Rad).

Statistical Data Analysis

Statistical differences among the potential progeny of mouse crosses were determined using the Fisher's Exact test. The student's t-test was used for estimating statistical differences between the weights of $F5^{L/L}$ $Tfpi^{+/-}$ mice and their littermates. Relative expression differences for *F3* alleles were estimated using the Wilcoxon rank-sum test. All the above statistical analyses were performed using the 'stats' package in R software. Kaplan-Meier survival curves with log-rank test to estimate significant differences in mouse survival as well as significance for putative suppressors identified by exome sequencing were performed using the 'survival' package in R(62).

Acknowledgements

This research was supported by NIH grants P01-HL057346 (to D. Ginsburg) and R15-HL133907 (to R. Westrick). R.J. Westrick was supported by the Oakland University Research Excellence Fund, and American Heart Association Predoctoral Fellowship, Innovative Research, and Scientist Development grants. K. Tomberg was an International Fulbright Science and Technology Fellow and the recipient of an American Heart Association Predoctoral Fellowship Grant. M.A. Brake and A. J. Johnston were recipients of American Heart Association Undergraduate Fellowship. D. Ginsburg is a member of the University of Michigan Cancer Center. Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under Award Number P30CA046592 by the use of the following Cancer Center Shared Resource(s): Transgenic Animal Models. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We gratefully acknowledge expertise of the Transgenic Animal Model Core staff of the University of Michigan's Biomedical Research Core Facilities for assistance with this study. D. Ginsburg in an Investigator of the Howard Hughes Medical Institute.

Figure legends

Figure 1: ***F8* deficient thrombosuppression and design of the Leiden ENU mutagenesis screen. A.** The mating scheme and observed distributions of the $F5^{L/+}$ $Tfpi^{+/-}$ *F8* deficiency rescue experiments. *F8 X* results in suppression of the $F5^{L/L}$ $Tfpi^{+/-}$ phenotype. **B.** The mating scheme and observed distribution of the Leiden screen. $F5^{L/L}$ $Tfpi^{+/+}$ male mice were mutagenized with either 1 x 150mg/kg or 3 x 90 mg/kg ENU and bred with non-mutagenized $F5^{L/+}$ $Tfpi^{+/-}$ females. Sixteen and 83 $F5^{L/L}$ $Tfpi^{+/-}$ progeny, respectively were observed in each of the dosing regimens, with over twice the rate of $F5^{L/L}$ $Tfpi^{+/-}$ survivors in the progeny of the 3 x 90 mg/kg treated mice. **C.** There was no significant difference in survival between male and

female $F5^{L/L}$ $Tfpi^{+/-}$ putative suppressor mice, (p=0.384). **D.** $F5^{L/L}$ $Tfpi^{+/-}$ putative suppressor mice were significantly smaller than their non-$F5^{L/L}$ $Tfpi^{+/-}$ littermates. **E.** $F5^{L/L}$ $Tfpi^{+/-}$ putative suppressors were smaller than their littermates of other genotypes ($p<8.8 \times 10^{-12}$ for B6, and $p=2.2 \times 10^{-16}$ for mixed B6-129S1) regardless of whether they were on the pure B6 or mixed B6-129S1 genetic backgrounds (p=0.327 between B6 and mixed backgrounds).

Figure 2: **The *MF5L6* suppressor locus maps to Chr3. A.** Linkage analysis for the *MF5L6* line. Alternating red and black is used to highlight the chromosomes. Chrs 1 and 2 were excluded from further analysis since they contain the *F5* and *Tfpi* genes, whose segregation was restricted by required genotypes at these loci. The Chr3 peak had the highest LOD score in the Chr3 subregion:117.3-124.8Mb (maximum LOD=4.49, 1 LOD interval, significance threshold of LOD >3.3(63)). **B.** The Chr3 candidate interval (Chr3:117.3-124.8 Mb) contains 43 refseq annotated genes, including *F3*. **C.** The mating scheme and observed distribution of offspring to test *F3* deficiency as a suppressor of $F5^{L/L}$ $Tfpi^{+/-}$. $F3^{+/-}$ results in incompletely penetrant suppression of the $F5^{L/L}$ $Tfpi^{+/-}$ phenotype.

Figure 3: **The *Actr2* R258G ENU-induced mutation is a potential thrombosis suppressor gene. A.** Kaplan-Meier survival plot for $F5^{L/L}$ $Tfpi^{+/-}$ mice with and without the $Actr2^{G}$ mutation. $F5^{L/L}$ $Tfpi^{+/-}$ $Actr2^{+/G}$ mice exhibit significantly better survival than $F5^{L/L}$ $Tfpi^{+/-}$ $Actr2^{+/+}$ littermates (n=19 for $Actr2^{+/G}$, n=12 for $Actr2^{+/+}$, n=31 total). **B.** ARP2 amino acid R258 is highly conserved in animals, plants and fungi.

Figure S1: **Relative expression analysis of *F3* alleles**. Expression differences between B6 and 129S1 alleles were estimated at three different SNPs: 1) rs30268372, 2) rs30269285, and 3) 30269288 using both forward (Fwd) and reverse (Rev) sequences of cDNA extracted from lung (red), liver (blue), and whole brain (green) tissues of adult mice. No significant differences were observed between the alleles in any of the tested tissues.

Figure S2**: Strategy to generate independent *Actr2* knockin ($Actr2^{G}$) allele by CRISPR/Cas9 technology. A.** Schematic representation of the pX459 vector and the designed guide sequence inserts. Image is an adaptation from Ran *et al*. (58). The guide oligos contain overhangs for ligation into the pair of BbsI sites (yellow highlight) in pX459. Digestion of pX459 with BbsI allows the replacement of the Type II restriction sites (red arrows) with direct insertion of annealed oligos. Likewise, a G-C base pair (blue text in sgRNA insert sequences) is added at the 5′ end of the guide sequence for U6 transcription. Color coding of the sequence correlates with positions within the guide RNA expression cassette located in the pX459 vector (Green text is 3' end of U6 promoter; bright blue text on yellow highlight are BbsI sites; black is chimeric guide RNA backbone and blue text is tracrRNA sequence. **B.** sgRNA and HDR donor design for targeting the mouse *Actr2* locus for R258G incorporation. *Actr2* gene is shown in the forward direction. Black and blue filled boxes indicate exons and the open reading frame, respectively. Grey text indicates the in-frame amino acid sequence, blue and red text indicate the C->G nucleotide change resulting in the R258G substitution and green text indicates synonymous mutations to block subsequent CRISPR binding following sequence replacement. Numbers in orange indicate the position within the 161 bp HDR template that is depicted in the figure. The blue and purple bars indicate the sgRNA and PAM respectively for sgRNA #1 and #2 and red

arrowheads indicate the DSB position within *Actr2* for each guide. *Actr2* locus scale bar, 1 kb (intronic regions reduced for fit). **C.** Schematic representation of *Actr2* primer locations. Note reverse orientation. **D.** Genotyping of heterogenous N2a cell populations post-puromycin selection. PCR was performed as described in Materials and Methods and depicted in Figure S2C. (Wt and Mut indicates competitive PCR with primers ACTR_OF/R and WT-R or MUT-R, respectively; NTC=no template control). In post-selected heterogeneous N2a cells, the 318 bp band indicating the presence of the $Actr2^{G}$ allele within the heterogenous cell population was observed when PCR was performed with the Mutant-specific primer (Mut) on gDNA isolated from cells co-transfected with sgRNA #1 + DONOR.

Figure S3: ***Actr2*** **hemizygosity is incompatible with survival. A.** *Actr2* R258G genotyping of three injected 60-cell stage mouse blastocysts and tail DNA from B6 control mouse. This is the same PCR scheme as Figure S2D. As expected, no mutant amplification is seen in B6 tail DNA sample. Amplification of the mutant allele is present in blastocysts #1 and #6 but not in #13. **B.** Surveyor assay for INDEL detection from products amplified with ACTR2_OF/R (C=cut; U=uncut). **C.** Sanger Sequencing chromatograms from the mouse blastocysts and B6 control. **D.** INDEL frequency for the three blastocysts determined by decomposition using TIDE software(64). Results indicate that all 3 blastocysts contain indels. Integration of the HDR donor template carrying the *Actr2* mutation and synonymous PAM change within the blastocysts were detected by both PCR (Fig. S3A and B) and Sanger sequencing (Fig. S3D). While integration of the HDR donor template is evident in blastocysts #1 and #6, they also contained multiple indels(64) (Fig. S3B-D). Genetic mosaicism can occur with high frequency in founder mice derived using CRISPR/Cas9 genome editing(65). In our targeting strategy, the degree of observed mosaicism was likely due to persistent Cas9 expression.

References

1. Silverstein MD, *et al.* (1998) Trends in the incidence of deep vein thrombosis and pulmonary embolism: a 25-year population-based study. *Arch Intern Med* 158(6):585-593.
2. Souto JC, *et al.* (2000) Genetic determinants of hemostasis phenotypes in Spanish families. *Circulation* 101(13):1546-1551.
3. Tregouet DA, *et al.* (2016) Is there still room for additional common susceptibility alleles for venous thromboembolism? *J Thromb Haemost*.
4. Dentali F, *et al.* (2012) Non-O blood type is the commonest genetic risk factor for VTE: results from a meta-analysis of the literature. *Semin Thromb Hemost* 38(5):535-548.
5. Rosendaal FR & Reitsma PH (2009) Genetics of venous thrombosis. *J Thromb Haemost* 7 Suppl 1:301-304.
6. Dahlback B (2008) Advances in understanding pathogenic mechanisms of thrombophilic disorders. *Blood* 112(1):19-27.
7. Lijfering WM, Rosendaal FR, & Cannegieter SC (2010) Risk factors for venous thrombosis - current understanding from an epidemiological point of view. *Br J Haematol* 149(6):824-833.
8. Clark JS, Adler G, Salkic NN, & Ciechanowicz A (2013) Allele frequency distribution of 1691G >A F5 (which confers Factor V Leiden) across Europe, including Slavic populations. *J Appl Genet* 54(4):441-446.
9. Westrick RJ & Ginsburg D (2009) Modifier genes for disorders of thrombosis and hemostasis. *Journal of thrombosis and haemostasis : JTH* 7 Suppl 1:132-135.
10. Anonymous (2011) 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* 13(1):67-76.
11. De Stefano V, *et al.* (1999) The risk of recurrent deep venous thrombosis among heterozygous carriers of both factor V Leiden and the G20210A prothrombin mutation. *N Engl J Med* 341(11):801-806.
12. van Boven HH, *et al.* (1996) Factor V Leiden (FV R506Q) in families with inherited antithrombin deficiency. *Thromb Haemost* 75(3):417-421.
13. Cui J, *et al.* (2000) Spontaneous thrombosis in mice carrying the factor V Leiden mutation. *Blood* 96(13):4222-4226.
14. Eitzman DT, *et al.* (2002) Lethal perinatal thrombosis in mice resulting from the interaction of tissue factor pathway inhibitor deficiency and factor V Leiden. *Circulation* 105(18):2139-2142.
15. Cordes SP (2005) N-ethyl-N-nitrosourea mutagenesis: boarding the mouse mutant express. *Microbiol Mol Biol Rev* 69(3):426-439.
16. Moresco EM, Li X, & Beutler B (2013) Going forward with genetics: recent technological advances and forward genetics in mice. *Am J Pathol* 182(5):1462-1473.
17. Bull KR, *et al.* (2013) Unlocking the bottleneck in forward genetics using whole-genome sequencing and identity by descent to isolate causative mutations. *PLoS Genet* 9(1):e1003219.

18. Buchovecky CM, *et al.* (2013) A suppressor screen in Mecp2 mutant mice implicates cholesterol metabolism in Rett syndrome. *Nat Genet* 45(9):1013-1020.
19. Tchekneva EE, *et al.* (2007) A sensitized screen of N-ethyl-N-nitrosourea-mutagenized mice identifies dominant mutants predisposed to diabetic nephropathy. *J Am Soc Nephrol* 18(1):103-112.
20. Matera I, *et al.* (2008) A sensitized mutagenesis screen identifies Gli3 as a modifier of Sox10 neurocristopathy. *Hum Mol Genet* 17(14):2118-2131.
21. Carpinelli MR, *et al.* (2004) Suppressor screen in Mpl-/- mice: c-Myb mutation causes supraphysiological production of platelets in the absence of thrombopoietin signaling. *Proc Natl Acad Sci U S A* 101(17):6553-6558.
22. Bank I, *et al.* (2005) Elevated levels of FVIII:C within families are associated with an increased risk for venous and arterial thrombosis. *J Thromb Haemost* 3(1):79-84.
23. Nolan PM, *et al.* (2000) A systematic, genome-wide, phenotype-driven mutagenesis programme for gene function studies in the mouse. *Nat Genet* 25(4):440-443.
24. Justice MJ, *et al.* (2000) Effects of ENU dosage on mouse strains. *Mamm Genome* 11(7):484-488.
25. Weber JS, Salinger A, & Justice MJ (2000) Optimal N-ethyl-N-nitrosourea (ENU) doses for inbred mouse strains. *Genesis* 26(4):230-233.
26. Davis AP & Justice MJ (1998) An Oak Ridge legacy: the specific locus test and its role in mouse mutagenesis. *Genetics* 148(1):7-12.
27. Cordes SP (2005) N-ethyl-N-nitrosourea mutagenesis: boarding the mouse mutant express. *Microbiol Mol Biol Rev* 69(3):426-439.
28. Lusis AJ (2012) Genetics of atherosclerosis. *Trends Genet* 28(6):267-275.
29. Yoo YJ & Mendell NR (2008) The power and robustness of maximum LOD score statistics. *Ann Hum Genet* 72(Pt 4):566-574.
30. Kyrle PA, *et al.* (2004) The risk of recurrent venous thromboembolism in men and women. *N Engl J Med* 350(25):2558-2563.
31. Roach RE, *et al.* (2015) Sex difference in the risk of recurrent venous thrombosis: a detailed analysis in four European cohorts. *J Thromb Haemost* 13(10):1815-1822.
32. Vandenbroucke JP, *et al.* (2001) Oral contraceptives and the risk of venous thrombosis. *N Engl J Med* 344(20):1527-1535.
33. Justice MJ, Noveroske JK, Weber JS, Zheng B, & Bradley A (1999) Mouse ENU mutagenesis. *Hum Mol Genet* 8(10):1955-1963.
34. Rich JT, *et al.* (2010) A practical guide to understanding Kaplan-Meier curves. *Otolaryngol Head Neck Surg* 143(3):331-336.
35. Rotty JD, Wu C, & Bear JE (2013) New insights into the regulation and cellular functions of the ARP2/3 complex. *Nat Rev Mol Cell Biol* 14(1):7-12.
36. Rottner K, Hanisch J, & Campellone KG (2010) WASH, WHAMM and JMY: regulation of Arp2/3 complex and beyond. *Trends in cell biology* 20(11):650-661.
37. Veltman DM & Insall RH (2010) WASP family proteins: their evolution and its physiological implications. *Molecular biology of the cell* 21(16):2880-2893.
38. Falet H, *et al.* (2002) Importance of free actin filament barbed ends for Arp2/3 complex function in platelets and fibroblasts. *Proc Natl Acad Sci U S A* 99(26):16782-16787.
39. Vauti F, *et al.* (2007) Arp3 is required during preimplantation development of the mouse embryo. *FEBS Lett* 581(29):5691-5697.

40. Li Z, Kim ES, & Bearer EL (2002) Arp2/3 complex is required for actin polymerization during platelet shape change. *Blood* 99(12):4466-4474.
41. Koseoglu S*, et al.* (2015) VAMP-7 links granule exocytosis to actin reorganization during platelet activation. *Blood* 126(5):651-660.
42. Falet H*, et al.* (2002) Importance of free actin filament barbed ends for Arp2/3 complex function in platelets and fibroblasts. *Proc Natl Acad Sci U S A* 99(26):16782-16787.
43. Lek M*, et al.* (2016) Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536(7616):285-291.
44. McKusick VA (2007) Mendelian Inheritance in Man and its online version, OMIM. *American journal of human genetics* 80(4):588-604.
45. Beutler B (2017) MUTAGENETIX (TM).
46. Westrick RJ & Ginsburg D (2009) Modifier genes for disorders of thrombosis and hemostasis. *J Thromb Haemost* 7 Suppl 1:132-135.
47. Toomey JR, Kratzer KE, Lasky NM, Stanton JJ, & Broze GJ, Jr. (1996) Targeted disruption of the murine tissue factor gene results in embryonic lethality. *Blood* 88(5):1583-1587.
48. Huang ZF, Higuchi D, Lasky N, & Broze GJ, Jr. (1997) Tissue factor pathway inhibitor gene disruption produces intrauterine lethality in mice. *Blood* 90(3):944-951.
49. Bi L*, et al.* (1995) Targeted disruption of the mouse factor VIII gene produces a model of haemophilia A. *Nat Genet* 10(1):119-121.
50. Lange K*, et al.* (2013) Mendel: the Swiss army knife of genetic analysis programs. *Bioinformatics* 29(12):1568-1570.
51. Lander E & Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 11(3):241-247.
52. Tomberg K*, et al.* (2016) Spontaneous 8bp Deletion in Nbeal2 Recapitulates the Gray Platelet Syndrome in Mice. *PLoS One* 11(3):e0150852.
53. Mohlke KL*, et al.* (1999) Mvwf, a dominant modifier of murine von Willebrand factor, results from altered lineage-specific expression of a glycosyltransferase. *Cell* 96(1):111-120.
54. Li H & Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14):1754-1760.
55. Garcia-Alcalde F*, et al.* (2012) Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* 28(20):2678-2679.
56. DePristo MA*, et al.* (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43(5):491-498.
57. Wang K, Li M, & Hakonarson H (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 38(16):e164.
58. Ran FA*, et al.* (2013) Genome engineering using the CRISPR-Cas9 system. *Nature protocols* 8(11):2281-2308.
59. Brinster RL, Chen HY, Trumbauer ME, Yagle MK, & Palmiter RD (1985) Factors affecting the efficiency of introducing foreign DNA into mice by microinjecting eggs. *Proc Natl Acad Sci U S A* 82(13):4438-4442.
60. Sakurai T, Watanabe S, Kamiyoshi A, Sato M, & Shindo T (2014) A single blastocyst assay optimized for detecting CRISPR/Cas9 system-induced indel mutations in mice. *BMC Biotechnol* 14:69.

61. Yang H*, et al.* (2013) One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* 154(6):1370-1379.
62. Therneau TM, Grambsch, P.M. (2000) *Modeling survival data: extending the Cox model* (Springer-Verlag New York) 1 Ed pp XIV, 350.
63. Lander E & Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 11(3):241-247.
64. Brinkman EK, Chen T, Amendola M, & van Steensel B (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* 42(22):e168.
65. Singh P, Schimenti JC, & Bolcun-Filas E (2015) A mouse geneticist's practical guide to CRISPR applications. *Genetics* 199(1):1-15.

Figure 1

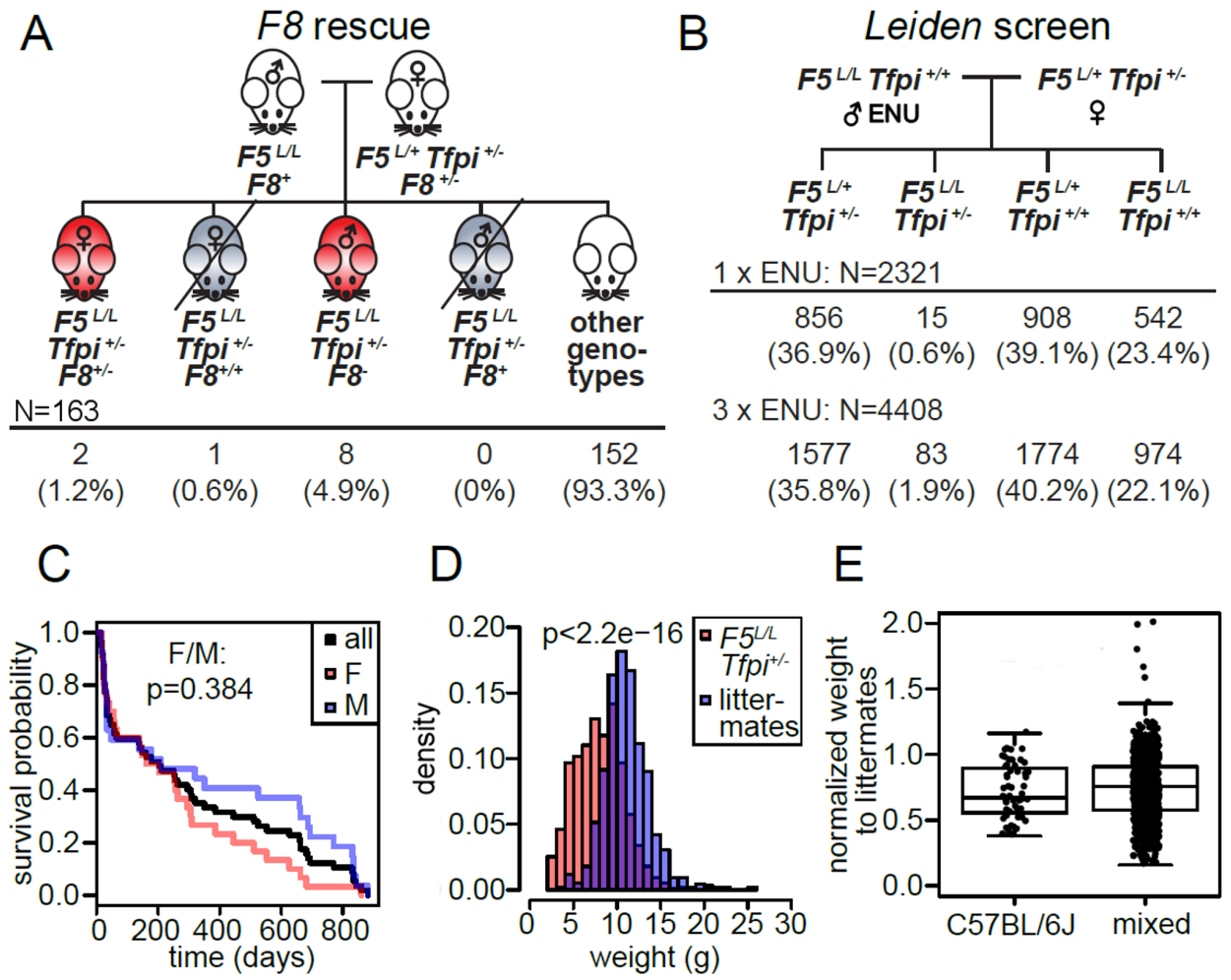

Figure 2

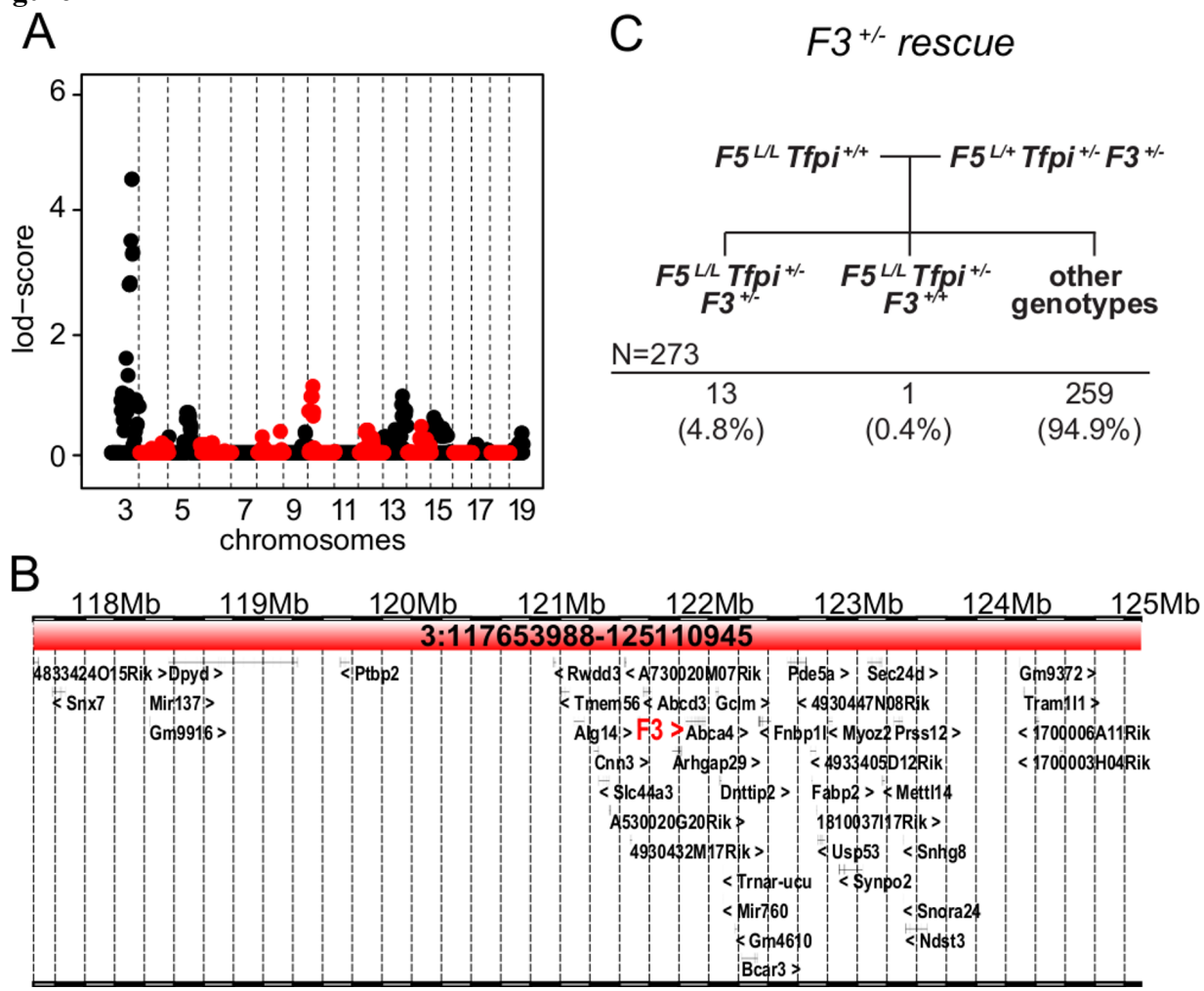

Figure 3

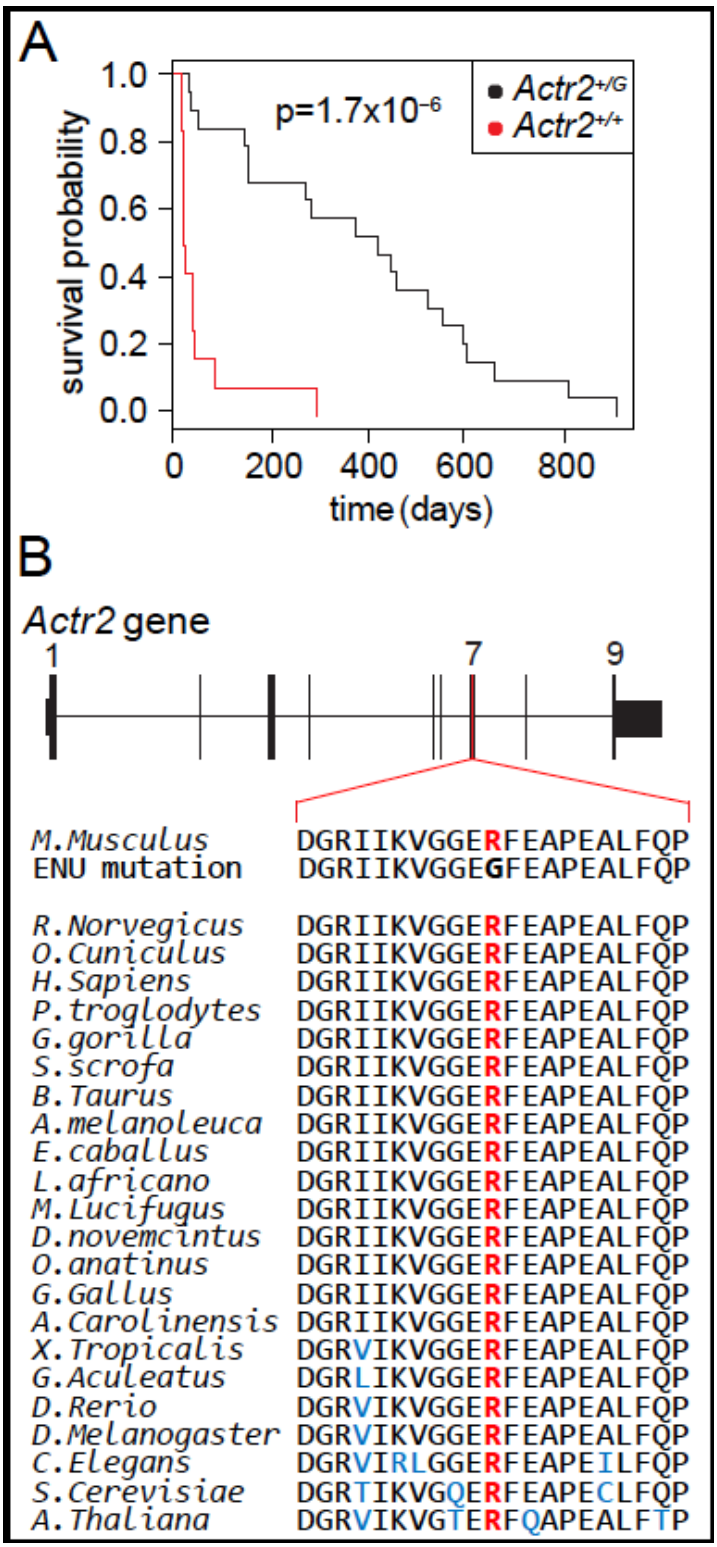

Figure S1

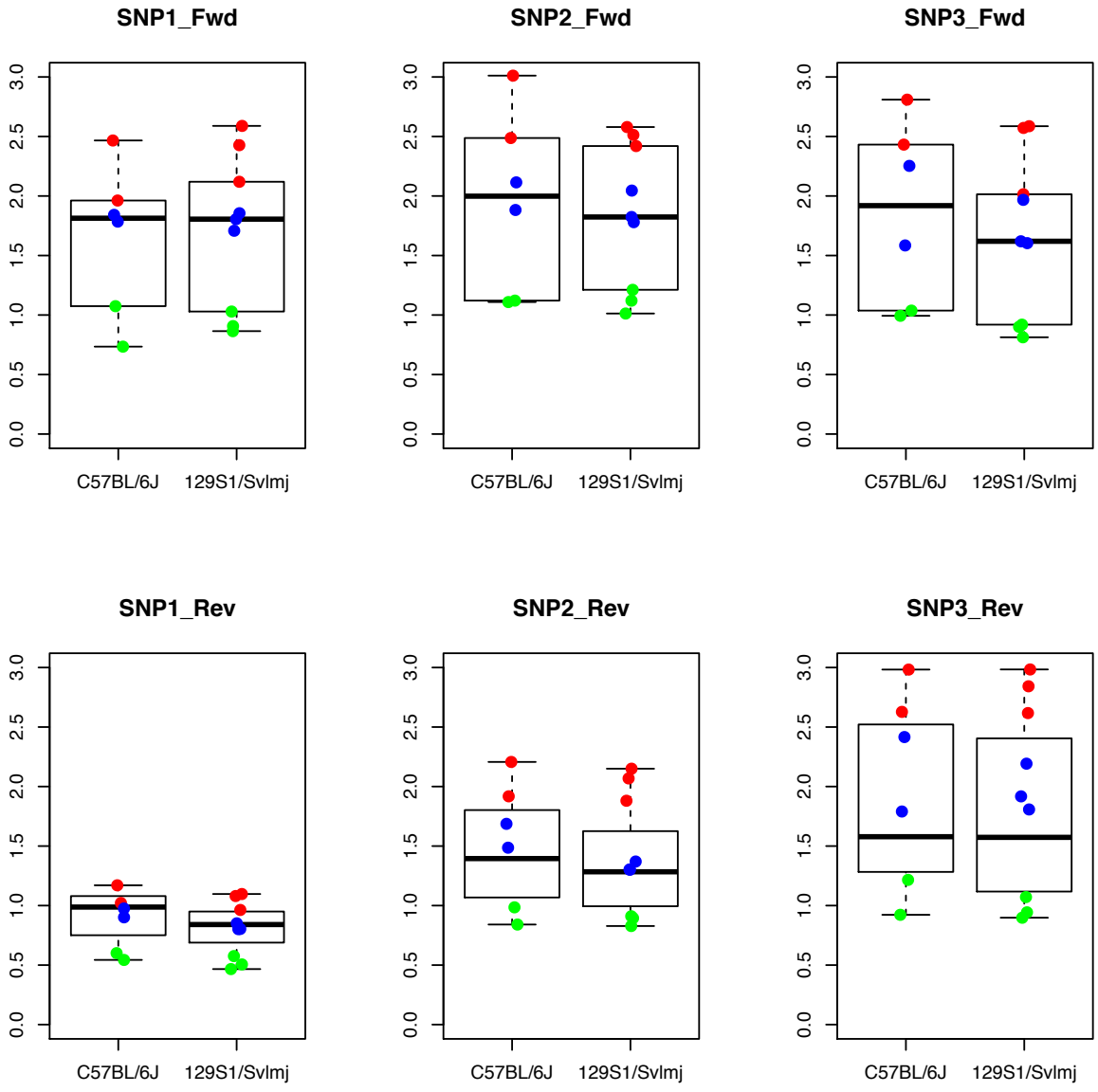

Figure S2

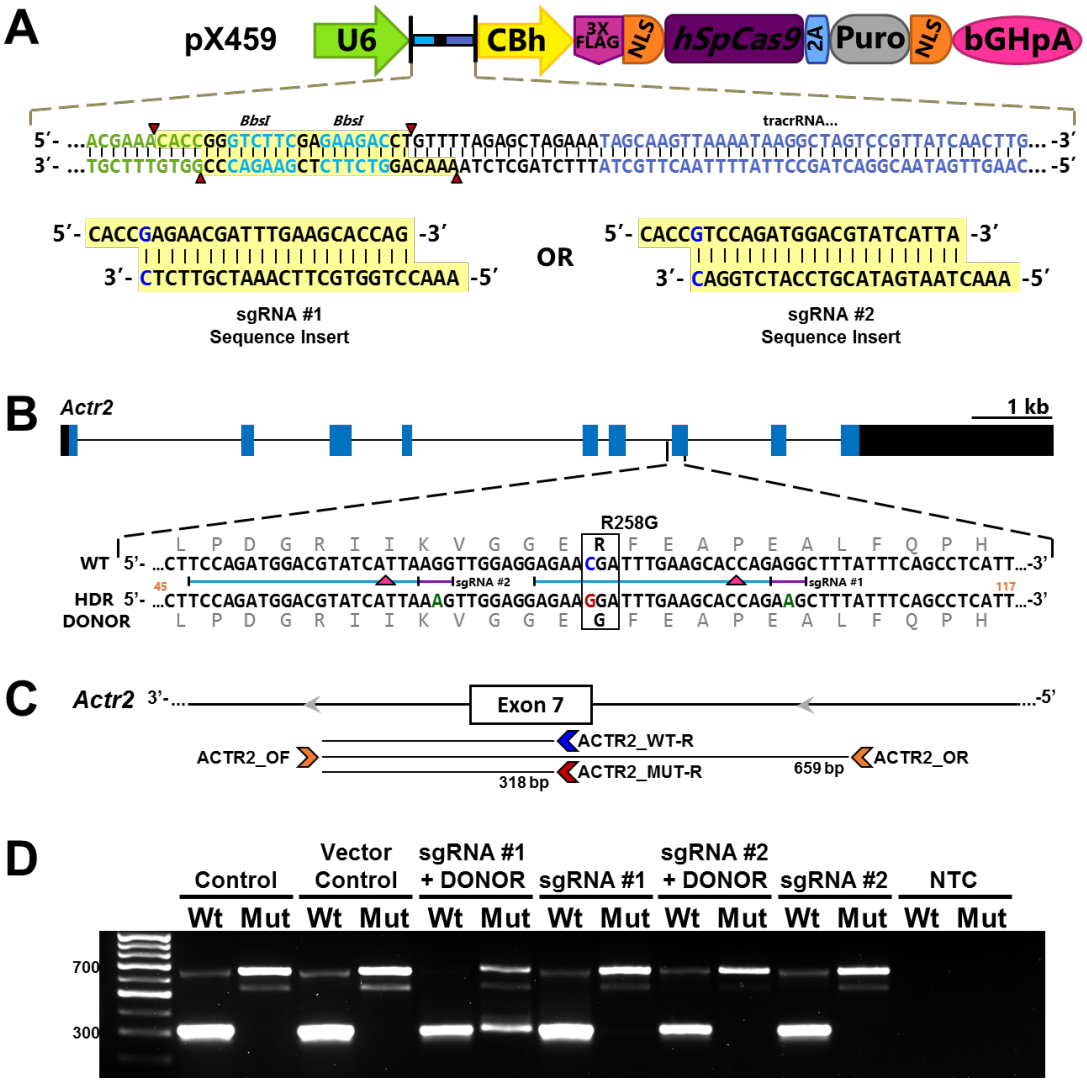

Figure S3

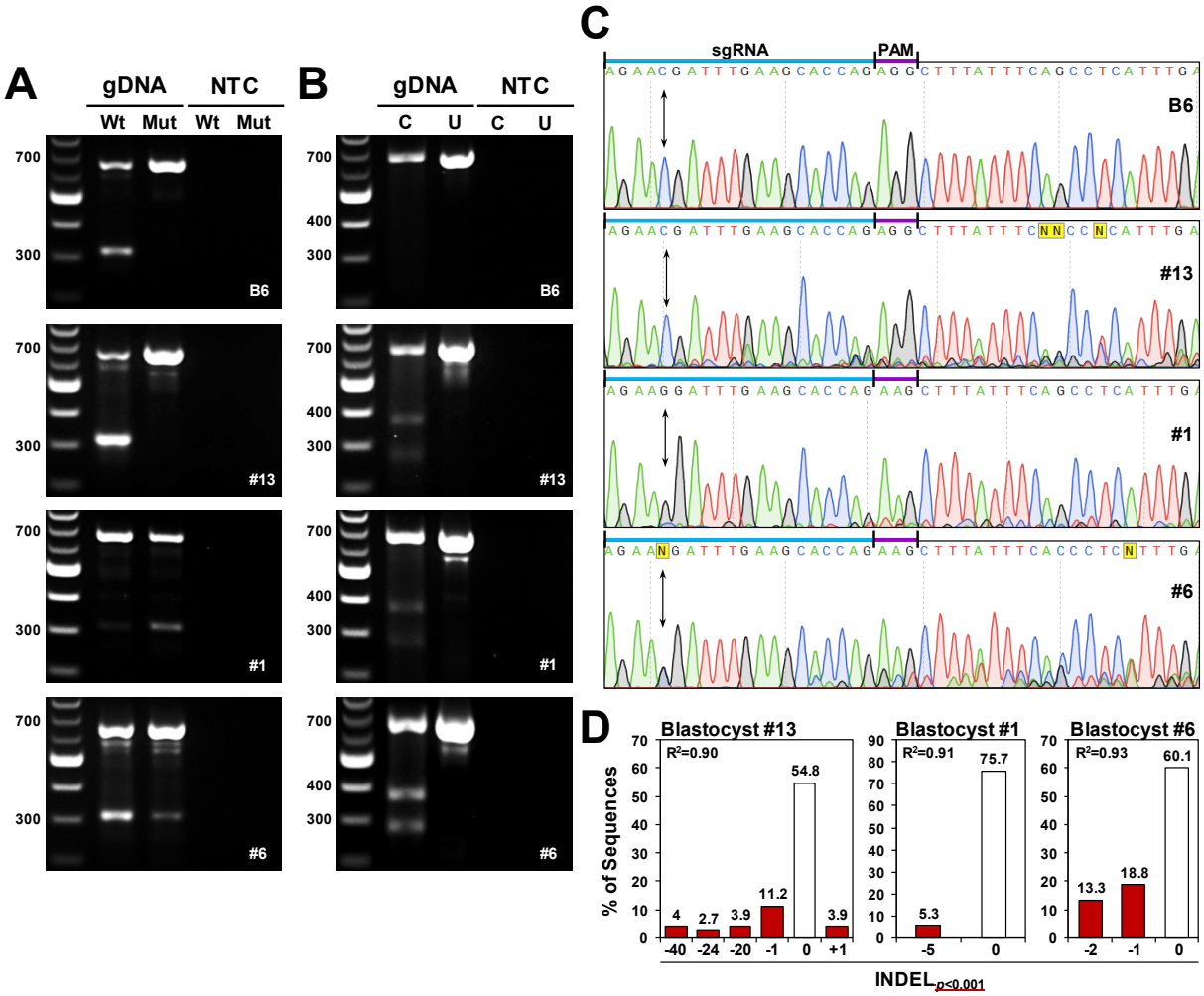

Table S1: Progeny Genotypes and Penetrance of Putative *MF5L* Thrombosuppressor Genes

ENU Lines	**Sex of G1**	**Total Progeny**	**Total *F5*$^{L/L}$ *Tfpi*$^{+/-}$ Mice**	***F5*$^{L/L}$ *Tfpi*$^{+/+}$ Littermates**	**Penetrance**
MF5L1	M	654	184	470	78.3%
MF5L2	F	14	1	13	15.4%
MF5L3	F	50	3	47	12.8%
MF5L4	M	3	1	2	100.0%
MF5L5	M	255	50	205	48.8%
MF5L6	M	1393	336	1057	63.6%
MF5L7	F	42	1	41	4.9%
MF5L8	M	543	132	411	64.2%
MF5L9	M	1127	264	863	61.2%
MF5L10	M	111	15	96	31.3%
MF5L11	M	459	121	338	71.6%
MF5L12	M	200	46	154	59.7%
MF5L13	M	115	13	102	25.5%
MF5L14	M	47	3	44	13.6%
MF5L15	F	40	3	37	16.2%
MF5L16	M	442	119	323	73.7%

Penetrance was calculated as total *F5*$^{L/L}$ *Tfpi*$^{+/-}$ divided by ½ of the number of *F5*$^{L/L}$ *Tfpi*$^{+/+}$ littermates.

Table S2: Test Cross Progeny from *F5*$^{L/+}$ *Tfpi*$^{+/-}$ X *F5*$^{L/L}$ Mating on the 129S1 Background and *F5*$^{L/+}$ X *F5*$^{L/+}$ Mating on the B6 Background

	129S1 Background *F5*$^{L/+}$ *Tfpi*$^{+/-}$ x *F5*$^{L/L}$				**B6 Background *F5*$^{L/+}$ x *F5*$^{L/+}$**		
F5 Genotype	L/+	L/+	L/L	L/L	+/+	L/+	L/L
Tfpi Genotype	+/+	+/-	+/+	+/-	+/+	+/+	+/+
Expected # of Conceptuses	25%	25%	25%	25%	25%	50%	25%
Observed, weaning age	43% (76)	42% (74)	**9.0% (27)**	0.0% (0)	27.1% (136)	55.8% (280)	**17.1% (86)**

A ~50% reduction of live weaning age *F5*$^{L/L}$ mice >N12 on the 129S1 background is observed compared to *F5*$^{L/L}$ mice >N12 on the B6 background (shown in bold).

Table S3: Exomic Variants from Whole Exome Sequencing

Chr	Pos	SNP	Ref	Alt	Variant type	ENU Line	Gene	Exon	AA change	Validation
2	70509665	rs33889947	G	C	nonsynonymous SNV	*MF5L1*	*Erich2*	exon2	C158S	chr2
7	14225894	---	T	C	nonsynonymous SNV	*MF5L1*	*Sult2a6*	exon5	Q238R	Not Verified
7	15940142	---	T	C	nonsynonymous SNV	*MF5L1*	*Gltscr2*	exon6	H254R	ENU
7	85754889	---	A	T	nonsynonymous SNV	*MF5L1*	*Vmn2r72*	exon1	D31E	Not Verified
8	108949251	---	A	G	nonsynonymous SNV	*MF5L1*	*Zfhx3*	exon9	Q2311R	Not Verified
12	110977486	---	G	A	nonsynonymous SNV	*MF5L1*	*Ankrd9*	exon3	T5I	Not Verified
17	25722876	---	T	A	nonsynonymous SNV	*MF5L1*	*Chtf18*	exon13	N541Y	ENU
3	99352190	---	A	G	nonsynonymous SNV	*MF5L5*	*Tbx15*	exon8	N459S	Not Verified
3	135228816	---	T	A	nonsynonymous SNV	*MF5L5*	*Cenpe*	exon13	D381E	ENU
6	35080128	---	G	A	nonsynonymous SNV	*MF5L5*	*Cnot4*	exon2	R3C	Not Verified
8	70913530	---	A	T	nonsynonymous SNV	*MF5L5*	*Map1s*	exon5	I360F	ENU
17	70657633	---	T	G	splicing	*MF5L5*	*Dlgap1*	exon4	c.1368+2T	ENU
1	82741945	---	C	A	nonsynonymous SNV	*MF5L6*	*Mff*	exon6	Q190K	de novo
2	67516594	---	A	G	nonsynonymous SNV	*MF5L6*	*Xirp2*	exon7	R3060G	chr2
2	76724952	---	G	A	nonsynonymous SNV	*MF5L6*	*Ttn*	exon167	R22243C	chr2
2	76939280	---	T	A	nonsynonymous SNV	*MF5L6*	*Ttn*	exon34	N2675I	chr2
2	88423385	---	A	T	nonsynonymous SNV	*MF5L6*	*Olfr1181*	exon1	F213L	chr2
2	111537791	---	A	T	nonsynonymous SNV	*MF5L6*	*Olfr1294*	exon1	L166Q	chr2
2	140120707	---	T	C	nonsynonymous SNV	*MF5L6*	*Esf1*	exon14	K815E	chr2
5	108650355	---	C	T	nonsynonymous SNV	*MF5L6*	*Dgkq*	exon18	R679H	Not ENU
7	79822260	---	A	C	nonsynonymous SNV	*MF5L6*	*Anpep*	exon20	D955E	ENU
7	101990583	---	A	T	stopgain SNV	*MF5L6*	*Numa1*	exon4	K47X	ENU
1	33746762	---	T	G	nonsynonymous SNV	*MF5L8*	*Bag2*	exon3	I160L	ENU
1	36163249	---	T	C	nonsynonymous SNV	*MF5L8*	*Uggt1*	exon29	Y1089C	ENU
1	40125203	---	A	G	nonsynonymous SNV	*MF5L8*	*Il1r2*	exon9	N410S	ENU
2	76549471	---	A	C	nonsynonymous SNV	*MF5L8*	*Osbpl6*	exon7	D135A	chr2
2	112407616	---	G	A	nonsynonymous SNV	*MF5L8*	*Katnbl1*	exon5	V152I	chr2
2	112630022	---	A	G	nonsynonymous SNV	*MF5L8*	*Aven*	exon4	T162A	chr2
2	153136757	---	G	A	nonsynonymous SNV	*MF5L8*	*Hck*	exon9	V276M	chr2

2	153225070	---	T	A	nonsynonymous SNV	*MF5L8*	*Tspyl3*	exon1	T83S	chr2
9	123712602	---	T	G	nonsynonymous SNV	*MF5L8*	*Lztfl1*	exon3	I51L	ENU
10	128290865	---	C	T	stopgain SNV	*MF5L8*	*Stat2*	exon23	Q820X	ENU
11	52145503	---	T	C	nonsynonymous SNV	*MF5L8*	*Olfr1373*	exon1	E9G	Not ENU
11	69129597	---	A	G	nonsynonymous SNV	*MF5L8*	*Aloxe3*	exon4	M156V	Not ENU
14	8169757	---	A	G	nonsynonymous SNV	*MF5L8*	*Pdhb*	exon7	S218P	ENU
15	89456795	---	G	T	stopgain SNV	*MF5L8*	*Mapk8ip2*	exon3	G148X	ENU
16	59554543	---	C	T	nonsynonymous SNV	*MF5L8*	*Crybg3*	exon1	R402H	Not ENU
17	12271353	---	T	A	nonsynonymous SNV	*MF5L8*	*Map3k4*	exon3	N397I	ENU
17	45416968	---	T	A	stopgain SNV	*MF5L8*	*Cdc5l*	exon7	K294X	ENU
19	39563826	---	C	T	nonsynonymous SNV	*MF5L8*	*Cyp2c39*	exon7	A321V	ENU
19	46065668	---	C	G	nonsynonymous SNV	*MF5L8*	*Pprc1*	exon6	I1150M	ENU
2	101696795	---	C	T	nonsynonymous SNV	*MF5L9*	*Traf6*	exon8	R297C	chr2
7	82868974	---	G	A	nonsynonymous SNV	*MF5L9*	*Mex3b*	exon2	G166R	ENU
9	21634876	---	T	C	nonsynonymous SNV	*MF5L9*	*Smarca4*	exon3	S117P	ENU
10	67538372	---	T	C	nonsynonymous SNV	*MF5L9*	*Egr2*	exon1	M51T	ENU
18	71327504	---	C	T	nonsynonymous SNV	*MF5L9*	*Dcc*	exon24	D1172N	de novo
19	56810315	---	G	T	stopgain SNV	*MF5L9*	*A630007-B06Rik*	exon2	S247X	ENU
2	30086662	---	A	G	nonsynonymous SNV	*MF5L11*	*Pkn3*	exon15	K572E	chr2
2	40874986	---	G	T	nonsynonymous SNV	*MF5L11*	*Lrp1b*	exon55	Q2943K	chr2
2	61804747	---	T	C	nonsynonymous SNV	*MF5L11*	*Tbr1*	exon1	S14P	chr2
2	144572561	---	G	A	nonsynonymous SNV	*MF5L11*	*Sec23b*	exon10	G398R	chr2
4	141581029	---	A	G	nonsynonymous SNV	*MF5L11*	*Fblim1*	exon8	I323T	ENU
5	123760656	---	T	A	stopgain SNV	*MF5L11*	*Kntc1*	exon8	C204X	ENU
5	136373331	---	T	C	nonsynonymous SNV	*MF5L11*	*Cux1*	exon5	K144E	ENU
9	123963447	---	G	T	nonsynonymous SNV	*MF5L11*	*Ccr1*	exon2	H349N	ENU
10	84958016	---	A	G	nonsynonymous SNV	*MF5L11*	*Ric8b*	exon4	S248G	ENU
11	57221033	---	A	T	nonsynonymous SNV	*MF5L11*	*Gria1*	exon7	T224S	ENU
11	101740781	---	T	G	nonsynonymo(1)us SNV	*MF5L11*	*Dhx8*	exon8	V400G	ENU
13	112368238	---	G	A	stopgain SNV	*MF5L11*	*Ankrd55*	exon9	W506X	ENU

14	32966414	---	A	G	nonsynonymous SNV	*MF5L11*	*Wdfy4*	exon56	W2921R	ENU
16	92605854	---	T	C	nonsynonymous SNV	*MF5L11*	*Runx1*	exon8	Y400C	ENU
5	86719746	---	T	A	nonsynonymous SNV	*MF5L12*	*Tmprss11e*	exon5	D155V	ENU
6	129517379	---	A	G	nonsynonymous SNV	*MF5L12*	*Tmem52b*	exon5	E182G	ENU
6	148237808	---	G	A	nonsynonymous SNV	*MF5L12*	*Tmtc1*	exon20	R939W	ENU
7	141620530	---	G	A	nonsynonymous SNV	*MF5L12*	*Ap2a2*	exon12	G504E	ENU
11	20077297	---	G	C	nonsynonymous SNV	*MF5L12*	*Actr2*	exon7	R258G	ENU
11	67921730	---	C	T	nonsynonymous SNV	*MF5L12*	*Usp43*	exon1	G107S	ENU
13	100285719	---	C	T	nonsynonymous SNV	*MF5L12*	*Naip7*	exon14	A1269T	Not Verified
6	36523684	---	A	G	nonsynonymous SNV	*MF5L16*	*Chrm2*	exon3	I159V	ENU
8	70259804	---	G	A	nonsynonymous SNV	*MF5L16*	*Sugp2*	exon9	R1023Q	ENU
10	77260815	---	T	C	nonsynonymous SNV	*MF5L16*	*Pofut2*	exon2	F125L	ENU
10	114800967	---	T	C	nonsynonymous SNV	*MF5L16*	*Trhde*	exon1	S112G	ENU
10	117278121	---	T	C	stoploss SNV	*MF5L16*	*Lyz2*	exon4	X149W	ENU
11	60710357	---	G	A	nonsynonymous SNV	*MF5L16*	*Llgl1*	exon16	R707H	Introduced
13	34896062	---	T	A	stopgain SNV	*MF5L16*	*Prpf4b*	exon11	L803X	Introduced
13	61568333	---	A	T	nonsynonymous SNV	*MF5L16*	*Cts3*	exon3	H71Q	ENU
13	90898831	---	G	A	nonsynonymous SNV	*MF5L16*	*Atp6ap1l*	exon4	P76S	ENU
13	94443934	---	G	A	nonsynonymous SNV	*MF5L16*	*Ap3b1*	exon9	A321T	ENU
15	6786636	---	C	T	stopgain SNV	*MF5L16*	*Rictor*	exon31	R1130X	ENU

Table S4: Whole Exome Sequencing Data of *MF5L* Lines

Line	**Whole Exome**		**Agilent Capture Region**				
	# of Reads	Mapped %	# of Reads	Mapped %	Mean Coverage	% bp Covered at >=6X	Mapping Quality (max 60)
MF5L1	97388058	91.06%	43177469	44.34%	73.88	>97%	34.42
MF5L5	110246719	93.07%	52837528	47.93%	89.9	>97%	34.36
MF5L6	133205717	86.03%	56688221	42.56%	93.93	>97%	33.89
MF5L8	128521513	96.48%	61950221	48.20%	105.88	>98%	34.69
MF5L9	109882856	94.18%	55469742	50.48%	95.56	>97%	34.67
MF5L11	110448223	96.38%	54907249	49.71%	93.22	>97%	35.08
MF5L12	105612822	94.27%	50684794	47.99%	87.28	>97%	34.68
MF5L16	115987369	90.33%	63608463	49.54%	110.83	>98%	34.26
Average	**113911660**	**92.73%**	**54915461**	**47.59%**	**93.81**	**>97%**	**34.51**

Table S5: Sequences of Oligonucleotides Used as HDR Donor Templates and for Cloning sgRNA Expression Vectors

Template Name	**Sequence (5' to 3')**
sgRNA #1 Forward	CACCGAGAACGATTTGAAGCACCAG
sgRNA #1 Reverse	AAACCTGGTGCTTCAAATCGTTCTC
sgRNA #1 HDR Donor	CAGAGAAGTTGTGGTTTTGCTGACACCTGGTTATTTCTTTCTAG CTTCCAGATGGACGTATCATTAAGGTTGGAGGAGAAGGATTTG AAGCACCAGAAGCTTTATTTCAGCCTCATTTGATCAATGTCGAG GGGGTTGGTGTTGCTGAACTGCTTTTTAAC
sgRNA #2 Forward	CACCGTCCAGATGGACGTATCATTA
sgRNA #2 Reverse	AAACTAATGATACGTCCATCTGGAC
sgRNA #2 HDR Donor	CAGAGAAGTTGTGGTTTTGCTGACACCTGGTTATTTCTTTCTAG CTTCCAGATGGACGTATCATTAAAGTTGGAGGAGAAGGATTTG AAGCACCAGAGGCTTTATTTCAGCCTCATTTGATCAATGTCGAG GGGGTTGGTGTTGCTGAACTGCTTTTTAAC

Changes in template are underlined

Table S6: Genotyping and Sanger Sequencing Primers

PRIMER NAME	PRIMER SEQ 5'->3'
F3_genotyping_F	ctcccatttcttttcctcctc
F3_genotyping_R	ggggcgtttgtaaatggcgg
F3-Neo	cctgactaggggaggagtag
F3-exon-F	tgcttctcgaccacagacac
F3-exon-R	ctgcttcctgggctattttg
upstreamF3_1F	gacacgccatctgtccagta
upstreamF3_1R	caaaaaggtgggcagctaag
upstreamF3_2F	agcagctcctgcaactcact
upstreamF3_2R	gcacagaggaagagcaaagg
upstreamF3_3F	cacaggggcctttattttga
upstreamF3_3R	aaagtagggcaggggaaaaa
upstreamF3_4F	accatctttgaagcccagaa
upstreamF3_4R	aggatggagcagaactgagg
upstreamF3_5F	ctgtcctgggaaacctgtgt
upstreamF3_5R	catgcaccactgcacctatc
upstreamF3_6F	ccaggacagcctcgaactta
upstreamF3_6R	agaaaatggctgctgtgctt
upstreamF3_7F	tggcctagcaactgtattttga
upstreamF3_7R	cagaagctgctcagtcatgg
upstreamF3_8F	gtccttttcctgggaagaca
upstreamF3_8R	cagtttacaagcacccaggag
upstreamF3_9F	gcttcagcgacaagagttca
upstreamF3_9R	actcccaactgagcaaagga
upstreamF3_10F	tcttcacgcatgtctgcttt
upstreamF3_10R	tgctttgtacaatcttccttcc
upstreamF3_11F	tgagtgggacgacagcttag
upstreamF3_11R	cacttgcaagctttgggttt
upstreamF3_12F	tgtcgagcaaatgctaccag
upstreamF3_12R	gcagtggctagcagatcattc
upstreamF3_13F	tctcaggcttcatgttgcag
upstreamF3_13R	cccctcctgtaggaaactcc
F3gene_1F	ggtctccgcagtacctggat
F3gene_1R	ttctcaggaccaatgccact
F3gene_2F	gctcctgtagcgtagccaac
F3gene_2R	cttcaagggcccaacatcta
F3gene_3F	gccctgaggatttgaatgaa
F3gene_3R	tgtcacatggtgggatgcta
F3gene_4F	tcaggcaagacagagtgcat
F3gene_4R	catactgcaatccgtggaaa
F3gene_5F	acgtgtgtgggggactagc
F3gene_5R	cgctttctctggaatgccta
F3gene_6F	cacaccctctgctcttgaca

F3gene_6R	tgtaggatggcctggaactc
F3gene_7F	gccaggttaaaaccaaagca
F3gene_7R	cactgcttcagggcagtgta
F3gene_8F	cactgtggtcactgtgttgct
F3gene_8R	gaaaccaaaagcttgccaaa
F3gene_9F	ccaatgccctttctggtta
F3gene_9R	gcatgcatgaacacacacac
F3gene_10F	gacagctctcgggaacaagt
F3gene_10R	caagctgtgcagggattaca
F3gene_11F	tggtgatgcaggtcagttgt
F3gene_11R	tgccttgactaatggcaatg
F3gene_12F	aaggtggtcaccattgaggt
F3gene_12R	tatggactggatggacagca
F3gene_13F	tcacactgactgctggtggt
F3gene_13R	gggctctgggtgaagtcata
F3gene_14F	tgctgtccatccagtccata
F3gene_14R	acattcagcaggggagtcac
F3gene_15F	tgggtcaaacaaaacactgc
F3gene_15R	aaagaacccagcacctcctt
F3gene_16F	tttgtgcctcttctgtgtgg
F3gene_16R	tctgcttagcgctcttctcc
F3gene_17F	attctgctgggctctttgaa
F3gene_17R	gagctgggtttgtttgcttc
F3gene_18F	ggagatctggaactcgcttg
F3gene_18R	tgtctgtggtcgagaagcac
F3gene_19F	tcggaggctcagactttgtt
F3gene_19R	taaaaactttggggcgtttg
F3gene_20F	tcccgtttcttttcctcctt
F3gene_20R	cccctggtctgatgaaagaa
F3gene_21F	cacacacaccaaggagatgc
F3gene_21R	aggggacagatggggattac
F3gene_22F	gtgtgtgagcctgccatcta
F3gene_22R	acacatcccacacccaatct
F3gene_23F	ggatgaagggcaattgagaa
F3gene_23R	atgcattagaggctgggaag
F3gene_24F	agattgggtgtgggatgtgt
F3gene_24R	tggtgacggtcttgtagctg
F3gene_25F	cctggtagccatcactcaca
F3gene_25R	gcatgctgtggagaatcaaa
F3gene_26F	catctgcaaggaagggtctc
F3gene_26R	ggggtccccaatatgaagat
F3gene_27F	caagcacgggaaaggtaaga
F3gene_27R	attgacgcacgagggattag
F3gene_28F	gtatgtgcttgcgtgtgtga
F3gene_28R	ggaagtgaccaagggaacaa

F3gene_29F	caaaatagcccaggaagcag
F3gene_29R	gctactgccccttagtcgt
F3gene_30F	ttgttcccttggtcacttcc
F3gene_30R	atgccccttggtctctttct
F3gene_31F	tagctatggcctggctctgt
F3gene_31R	tgatggtggagacgaagaga
F3gene_32F	ttctgccttcttgcctctgt
F3gene_32R	accactgctcccacaatgat
F3gene_33F	ccccagccaactactgtctc
F3gene_33R	atgttgcacagttcccatca
F3gene_34F	cgagcctccatgttgacttt
F3gene_34R	aatcacaaagatgccccaag
F3gene_35F	ccagctaacgctttgattcc
F3gene_35R	ttgtctcaattcccaatcacc
MFF_OF	cactcattgctgggtccttt
MFF_OR	atttccaagtgcaaccaagc
MFF_IF	ccctctgctcggattgatac
MFF_IR	tatgcaacaaagtggcaagg
DGKQ_OF	tgtccaaaactgtgccagac
DGKQ_OR	ccacacaggttccacctttt
DGKQ_IF	ccacaggcttcagtcaacaa
DGKQ_IR	acaggtgggcttagtcatcg
ANPEP_OF	tagcttcagagctgggcttc
ANPEP_OR	gggctgtggtttcacaactt
ANPEP_IF	ctccagaggctggagacttc
ANPEP_IR	ggtgagcacttaaccccaaa
NUMA1_OF	tcccaaacattttgccattt
NUMA1_OR	ttttcttgcaagggaaagga
NUMA1_IF	cttacccgccacacattttc
NUMA1_IR	ctggacctgacacggactct
BAG2_OF	ctgtgtctgccaacactgga
BAG2_OR	gttgctgacgtgggaagttt
BAG2_IF	ccggtgaatttgaaggctaa
BAG2_IR	gactgccaaccgtctgatg
UGGT1_OF	tccctagcactgcttcctgt
UGGT1_OR	gcagaaggcttggcttattg
UGGT1_IF	cgggaaggcatctgaataaa
UGGT1_IR	gacgctgagactgcatcaag
IL1R2_OF	ccacctaacccaagcctcta
IL1R2_OR	ggctgctatggcttgttctc
IL1R2_IF	gtcaacctatggtgccctgt
IL1R2_IR	cctccacattttctcccaga
LZTFL1_OF	tgagtgctcctcaaggaagg
LZTFL1_OR	cagaaagtgggggagttaagtg
LZTFL1_IF	agtgactgtgccttgctgtt

LZTFL1_IR	tgtcatgatgtcggtctcttg
STAT2_OF	tgtcccattgtctgtccttg
STAT2_OR	gcccttgcatttcctatcaa
STAT2_IF	gaccaggagttgccattgat
STAT2_IR	aggtcctcaggcaaatctga
OLFR1373_OF	cagggtgcataatggttgtg
OLFR1373_OR	acaccagggccaagaagt
OLFR1373_IF	caccttccaaagctgatggt
OLFR1373_IR	gggggagggtatagggaact
ALOXE3_OF	aatcggtgctgggatctatg
ALOXE3_OR	aagtctcaaccctgcccttt
ALOXE3_IF	tgaggcttagggatggctta
ALOXE3_IR	catctcaacacacggtggtc
PDHB_OF	actggtcttgaatgggcaac
PDHB_OR	ggggcatctagtgaggctta
PDHB_IF	ggcagctatggcctgtctta
PDHB_IR	ctgcatacctgcacatttgg
MAPK8IP2_OF	gcagccacaccctatttgtt
MAPK8IP2_OR	tacttcatggcgctcctctt
MAPK8IP2_IF	acaggcacttgctggagact
MAPK8IP2_IR	cagagcagggagttgggtta
CRYBG3_OF	tgagtcctggaagtctgcaa
CRYBG3_OR	gtctctcctgtttcccgaca
CRYBG3_IF	actggaggtcgttggttcac
CRYBG3_IR	tgaggcatttgatggagaca
MAP3K4_OF	cttcagtgctttgtccacga
MAP3K4_OR	cctcaggagacaaaccgtgt
MAP3K4_IF	tcctctgactcgagcctctc
MAP3K4_IR	agcaggtgaagcggataatg
CDC5L_OF	caagaactgccaccacttga
CDC5L_OR	gcctccatttatctttttctgc
CDC5L_IF	accgtgtttagtgccctcat
CDC5L_IR	tgcctgtgtgtaatctttttctg
CYP2C39_OF	gacaacagggcagatggagt
CYP2C39_OR	ctgccctctggaccataaag
CYP2C39_IF	aacactagtgaccttaaccaagga
CYP2C39_IR	cacggggtatgttgttaggg
PPRC1_OF	gacccaggagaacagaccaa
PPRC1_OR	ccctgcatcctctcttcatc
PPRC1_IF	tgttgcaaagctacctgctg
PPRC1_IR	aaaggaggcacagacgagaa
MEX3B_OF	cctggcttccaggttgtaaa
MEX3B_OR	gttgcgatagctggagaagg
MEX3B_IF	ggaggagcctgtctttgttg
MEX3B_IR	agatcaaagcccacgtctgt

SMARCA4_OF	tggtgagtgcctcagagcta
SMARCA4_OR	tgaaccccaggacctagtga
SMARCA4_IF	tctgtgtggtccccttctc
SMARCA4_IR	ttgctagcctccaggctcta
EGR2_OF	ggagggcaaaaggagatacc
EGR2_OR	ctagcccagtagcgcagagt
EGR2_IF	agttgggtctccaggttgtg
EGR2_IR	gcttcaaggaccaggagatg
DCC_OF	gaaggaaggcaacaggatga
DCC_OR	ctggggattcatctcagcat
DCC_IF	cttttctcaccccaaagcaa
DCC_IR	ggaaagacagccaggacaag
A630007B06RIK_OF	agtgccaaagtgtcccaaag
A630007B06RIK_OR	tcgtctgcttgcttctcttg
A630007B06RIK_IF	tttgggcagaaaatgtgcta
A630007B06RIK_IR	cagtcactcgatggtgagga
FBLIM1_OF	tggtccagtttgccacctat
FBLIM1_OR	gcacaatgggtagctggatt
FBLIM1_IF	ggctcgccacctatgtttt
FBLIM1_IR	acccctgtcgggaagagtag
KNTC1_OF	gccattgagaacacggactt
KNTC1_OR	tgatttatgggagggtgcat
KNTC1_IF	tcagccaagaaggtaagcaaa
KNTC1_IR	tcatcgagcctctagccttt
CUX1_OF	gacccttgatcaggagctg
CUX1_OR	ggcttgcctagaattcacca
CUX1_IF	ggcgacacatcagtctttga
CUX1_IR	gtgcagcgtctacacgacat
CCR1_OF	ggaatgccccattttgttta
CCR1_OR	tgctatgcagggatcatcag
CCR1_IF	gaccttccttggttgacacc
CCR1_IR	ctgctcagaagacccagtga
RIC8B_OF	gaacagaagaaccgggactg
RIC8B_OR	gcctgggagctactctcaaa
RIC8B_IF	ccctgaatggaatggagaga
RIC8B_IR	acaaatgcccaagtctgacc
GRIA1_OF	gaaggccaactgattttcca
GRIA1_OR	tggcatcacattttcatggt
GRIA1_IF	agctgatttgctggactggt
GRIA1_IR	gtcccacgtttgacttggat
DHX8_OF	cagtgctctcgttgtgcttt
DHX8_OR	cttcccttgccaccacag
DHX8_IF	acccagacagacccactcac
DHX8_IR	ccatggaacactgtctctgc
ANKRD55_OF	ccacctttgacagtgtcgtg

ANKRD55_OR	cagcccattcagggtagaaa
ANKRD55_IF	ccaccaatcagaacccagag
ANKRD55_IR	tggctgtagttcccgttttt
WDFY4_OF	cacacacacacatgcttgct
WDFY4_OR	ccccacacacacacctgtta
WDFY4_IF	ggcttgctcacccaataact
WDFY4_IR	gggcactttggtgtaccact
RUNX1_OF	agtttccctccgggattctt
RUNX1_OR	ggcagtctaggaagcctgtg
RUNX1_IF	gatggcgctcagctcagtag
RUNX1_IR	ctactctgccgtccatctcc
SULT2A6_OF	tgagggcttctggcaatagtc
SULT2A6_OR	gagctctctgttttcgccttca
SULT2A6_IF	gaggccttgggctttagcataca
SULT2A6_IR	tgagctggatctcgtcctcaa
GLTSCR2_OF	gaggtgagagtttggggtga
GLTSCR2_OR	aaactgtgagatgggctggt
GLTSCR2_IF	agggatgtccctctgatcct
GLTSCR2_IR	gtggttcttggcatggagtt
ANKRD9_OF	cacaagcgcatagagcagag
ANKRD9_OR	tagagttggggctggatttg
ANKRD9_IF	caggtagaaggcgaaggatg
ANKRD9_IR	ttgtgctgtgagcctgagtc
CHTF18_OF	ccaagtcaggaccctaccaa
CHTF18_OR	tgtattgagctgtcgggttg
CHTF18_IF	ctggcccactcatatcacct
CHTF18_IR	gttggctctgaggatgtgct
TBX15_OF	ctttttcaggggggcagtaca
TBX15_OR	ggccccattagaaggagaag
TBX15_IF	caacatggctgccttacaga
TBX15_IR	gagacgcgctcagtttctct
CENPE_OF	ggcaatcggcagattcat
CENPE_OR	tttaccgacagcctttaccg
CENPE_IF	gctactttcaagtttgctgcatt
CENPE_IR	agcctagagagttggagagtgg
CNOT4_OF	ggctctcaatccactgctct
CNOT4_OR	agccatcccattgatctcag
CNOT4_IF	aagcccattctcatcagtgc
CNOT4_IR	aacccttacatctctcttgcatc
MAP1S_OF	tgtcaacggctttacagtgc
MAP1S_OR	cagactgtcttggctgttgg
MAP1S_IF	caaccccaagtcgagcttct
MAP1S_IR	gcggccaggaaatagtacac
DLGAP1_OF	tgccctggtgttaaaagtcc
DLGAP1_OR	ttttgtcagtggggtcttcc

DLGAP1_IF	ccaaacgcttccattgttct
DLGAP1_IR	caaagagcccttcctctcct
TMPRSS11E_OF	tgaaagcctgcatgaaagaa
TMPRSS11E_OR	agacgagacagtgggtctgg
TMPRSS11E_IF	tcagcaaataaagggcaaga
TMPRSS11E_IR	tgtgctcgacagtgtctgct
TMEM52B_OF	ggatatcacctttgggtgtga
TMEM52B_OR	tgcatgctgtagttgaaggaa
TMEM52B_IF	ctcatgcacacagctccttg
TMEM52B_IR	ccttaatgttccatgctttgg
TMTC_OF	tatctgctgccaaagcactg
TMTC_OR	tgccattgaaccttatgctg
TMTC_IF	cagtgtgctctcctctgctg
TMTC_IR	gattccttaggggagccttg
AP2A2_OF	gagtggcctactggctgttc
AP2A2_OR	tgcacaagtggaacttggag
AP2A2_IF	atggtcttgccaaggcttta
AP2A2_IR	ctcatctgtcccatgtgctg
ACTR2_OF	ataaatgcaggctcccacaa
ACTR2_OR	ccggttggttggttttaag
ACTR2_IF	ctgtggcctatccgtgtctt
ACTR2_IR	agcttttgggggtcagtttt
ACTR2_MUT-R	atcattaaggrrggaggagaag
ACTR2_WT-R	atcattaaggttggaggagaac
USP43_OF	catgtcacctcgaaacaaagg
USP43_OR	tgcttaaccgcttcctgcta
USP43_IF	tggttcacacgcttttcaag
USP43_IR	ctacgatggcgatggagaag
LYZ2_OF	agcatccctcttgagcatcc
LYZ2_OR	cagaggtgtctgtgtggtga
LYZ2_IF	aggcagagcatcaaactgcc
LYZ2_IR	agggctgtgctgactgaca