Vaginal imperforation is caused by a point mutation in lipoma HMGIC fusion partner-like 2

Son Yong Karst, Heather E. Fairfield, Laura G. Reinholdt, David E. Bergstrom, and Leah Rae Donahue

Source of Support: This research was supported by NIH/NCRR grant RR01183 to the Mouse Mutant Resource (Leah Rae Donahue, PI) and Cancer Center Core Grant CA34196.

Mutation (allele) symbol: *vgim*

Mutation (allele) name: vaginal imperforation

Gene symbol: updated to *Lhfpl2* in October 2011

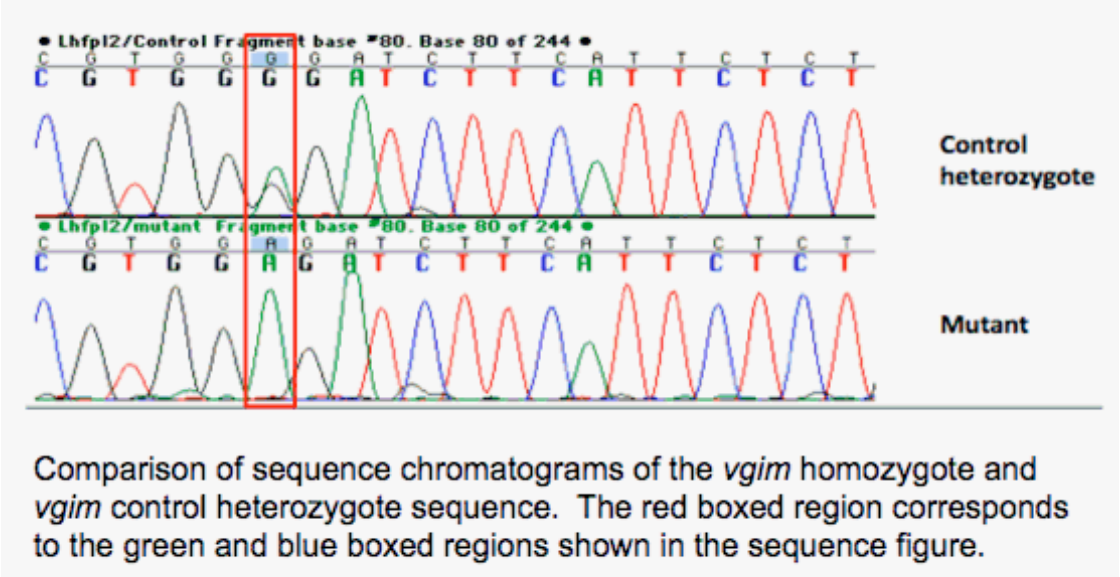
Strain of origin: B6.129S4-*Ttpa*^{tm1Far}/J

Current strain name: B6.Cg-*Lhfpl2*^{vgim}/GrsrJ

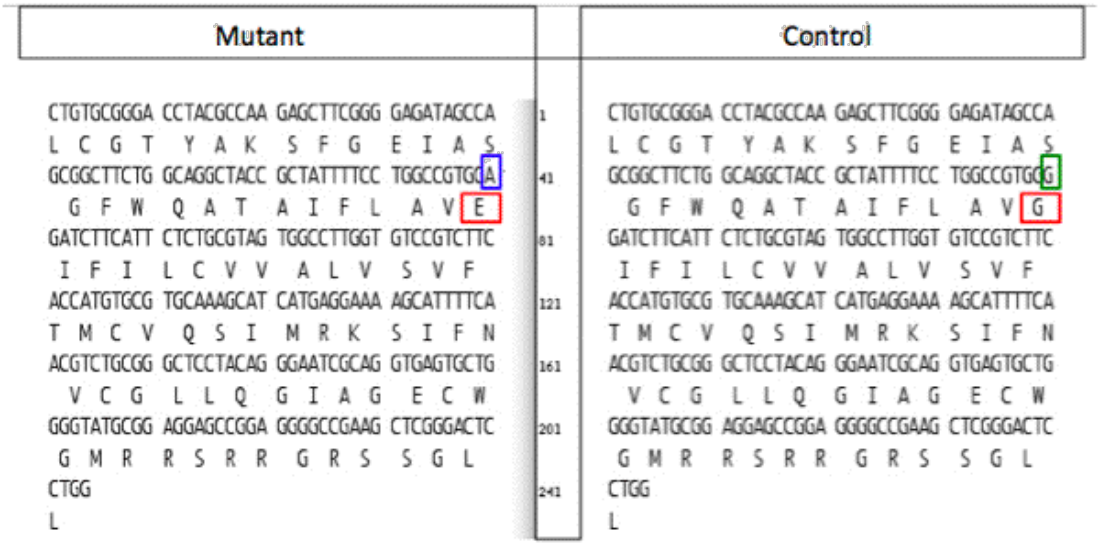
Stock #013716 (jaxmice.jax.org)

Phenotype categories: reproductive

The vaginal imperforation mutation (*vgim*) causes a developmental defect such that homozygous females have a complete closure of the vagina. Mapping data demonstrated linkage to Chromosome 13 and whole exome sequencing was used to identify candidate mutations in the mapped region. Briefly, genomic DNA was enriched for coding sequence by hybridization-based capture with probes representing 54 Mb of annotated coding sequence. The enriched DNA was then sequenced using the Illumina HiSeq high throughput sequencing platform[1]. A single nucleotide polymorphism was found at position 94,944,483 on chromosome 13 in lipoma HMGIC fusion partner-like 2 (*Lhfpl2*). Primers were generated that produce a 316 base pair product spanning the predicted mutation; *Lhfpl2* forward (CACTACCTGGGCATCCTCTG) and *Lhfpl2* reverse (CCAGCCCTCTTCAACAGATG). Sequence analysis of additional mutant and unaffected genomic DNA samples was used to confirm the presence of a single nucleotide transition from G to A at 116,733 in exon 3 (see chromatogram below) in affected DNA samples. This mutation would result in a predicted change of glycine to glutamic acid in amino acid 102 (see sequence figure below).



Comparison of sequence chromatograms of the *vgim* homozygote and *vgim* control heterozygote sequence. The red boxed region corresponds to the green and blue boxed regions shown in the sequence figure.



A portion of the protein coding region of *Lhfpl2*. The control DNA sequence and its amino acid translation are shown on the right, and the *vgim* mutant DNA and its translation on the left. A single nucleotide transition is enclosed by a blue box in the mutant sequence and a green box in the control sequence. The mutation is predicted to change amino acid 102 from glycine to glutamic acid.

[1]Exome Sequencing Protocol

The exome sequencing data were analyzed using tools and workflows provided by Genome Quest including processes for mapping (HS3), SNP calling and annotation of variants. Our analysis focused on novel variants, which were not positioned in repetitive sequence, had expected allele ratios (>0.95 for homozygous variants and >0.2 for heterozygous variants), and displayed sufficient locus coverage (at least 5X for homozygous variants and 10X for heterozygous variants) for effective mutation discovery. High priority was given to protein coding or splice variants within mapped regions, as well as unique variants that were not found in other exome data sets or in the Sanger Mouse Genomes Database. Following these analyses, re-sequencing of additional mutant and unaffected samples was performed to validate and determine the most likely causative mutation.