- Supplementary Information -

Targeted Deletion of the 9p21 Noncoding Coronary Artery Disease Risk Interval in Mice

Axel Visel, Yiwen Zhu, Dalit May, Veena Afzal, Elaine Gong, Catia Attanasio, Matthew J. Blow, Jonathan C. Cohen, Edward M. Rubin, and Len A. Pennacchio

Table of Contents

Supplementary Text

Reduced Survival in chr4$^{\Delta 70kb/\Delta 70kb}$ Mice

To assess a possible quantitative effect of the chr4$^{\Delta 70kb}$ allele on general viability, we compared the embryonic, postnatal and adult survival of chr4$^{+/\Delta 70kb}$ and chr4$^{\Delta 70kb/\Delta 70kb}$ mice to wild-type controls. We genotyped a total of 78 litters at embryonic stages between E9.5 and E15.5 to determine a possible effect of the deletion on prenatal survival. At stages E9.5 and E10.5, no significant deviation from expected Mendelian ratios was observed among surviving morphologically normal embryos (n = 191, $P = 0.29$, G-test). In contrast, among embryos collected between E11.5 and E15.5, chr4$^{+/\Delta 70kb}$ and chr4$^{\Delta 70kb/\Delta 70kb}$ genotypes were markedly depleted (113 wild-type : 200 heterozygous : 70 null embryos, P = 0.004), indicating that the chr4$^{\Delta 70kb}$ deletion negatively impacts on embryonic survival up to E15.5 (Suppl. Fig. 4a). Next, we tested a possible effect of the chr4$^{\Delta 70kb}$ deletion on postnatal viability. Since the genotype of pups from chr4$^{+/\Delta 70kb}$ × chr4$^{+/\Delta 70kb}$ crosses was generally not determined before weaning age (~3 weeks), we compared the survival from birth to weaning among 1,964 pups from wt × wt, chr4$^{+/\Delta 70kb}$ × chr4$^{+/\Delta 70kb}$ and chr4$^{\Delta 70kb/\Delta 70kb}$ × chr4$^{\Delta 70kb/\Delta 70kb}$ crosses. None of 345 pups from wt × wt crosses died before weaning. In contrast, 28 of 1,154 pups from heterozygous crosses (2.4%) and 24 of 465 pups from null crosses (5.2%) died before weaning (Suppl. Fig. 4b), suggesting that the chr4$^{\Delta 70kb}$ allele is associated with decreased survival of live-born pups to weaning age (P = 0.001 and P = 1.7×10^{-6}, respectively; two-tailed Fisher's Exact test). To determine if this decrease in viability also extends into adulthood, we monitored the survival of 16 chr4$^{\Delta 70kb/\Delta 70kb}$ mice that had survived to weaning and 16 wild-type controls for over one year under standard conditions with regular mouse chow fed *ad libitum* (Suppl. Fig. 4c). During this study period, 5 of the null mice (31%), but none of the wild-type controls died ($P = 0.016$, Kaplan-Meier test). The cause of death was usually not unambiguously determined and correlated only in some cases with the presence of tumors. Of note, in addition to expression changes in the heart and aorta, we also observed significant changes in expression of *Cdkn2a/b* in other tissues (Fig. 3c). Considering that these genes and the larger 9p21 locus have been implicated in several common forms of disease, the relative contribution of CAD-related phenotypes,

increased cancer incidence or other phenotypes remains to be determined. Taken together, these results indicate that the 70kb noncoding region containing the CAD risk interval is required for normal survival and suggest that molecular and physiological mechanisms negatively affected by the chr4$^{\Delta 70kb}$ deletion are not temporally restricted to a particular developmental stage, but rather persist throughout much of embryonic development and postnatal life in mice.

Effects of High-Fat Diet in chr4$^{\Delta 70kb/\Delta 70kb}$ Mice

Studying atherogenesis in mouse models represents a challenge since most mouse strains develop severe atherogenesis only upon genetic manipulation of their lipid metabolism [22,23]. However, the use of a high-fat, high-cholesterol diet ("Western" diet) can provide a model for fatty streak formation, an early stage of the atherogenic process, in the genetic background strain for the chr4$^{\Delta 70kb}$ deletion, 129Sv [21]. This strain is moderately susceptible to fatty lesion formation in this diet-induced model, allowing an initial assessment of genetic influences on early stages of atherogenic plaque formation [21]. To determine if the chr4$^{\Delta 70kb}$ deletion affects fatty streak formation, 40 chr4$^{\Delta 70kb}$ null mice and 40 wild-type controls were fed Western diet *ad libitum* for approximately 20 weeks. The diet caused marked aberrations in plasma lipid levels, namely a 2.5-fold reduction in triglycerides and a 2-fold increase in plasma cholesterol (Suppl. Fig. 6). However, no significant difference in this physiological response to the diet was observed between chr4$^{\Delta 70kb}$ null mice and controls. Following the Western diet regimen, formation of aortic fatty lesions was quantitated. Both in wild-type controls and in chr4$^{\Delta 70kb}$ null mice, the occurrence of lesions varied widely, from virtually absent to moderate levels. While a mild increase in median lesion size was observed in chr4$^{\Delta 70kb/\Delta 70kb}$ mice, the difference was overall not significant (Suppl. Fig. 7). Despite this lack of significant differences in plasma lipid levels and fatty lesion formation, we observed a substantially lowered tolerance for the high-fat diet in chr4$^{\Delta 70kb/\Delta 70kb}$ mice, resulting in increased mortality during the course of the feeding study (Suppl. Fig. 8). This was in sharp contrast to wild-type control animals in which cases of apparently diet-related deaths were observed only after several months of high-fat diet and at a lower frequency. The cause of death in chr4$^{\Delta 70kb/\Delta 70kb}$ animals could not be unambiguously determined, but was in most cases preceded by general signs of malaise over several days, including severe weight loss and

decreased activity. The etiological link between the high-fat diet and increased mortality in chr4$^{\Delta 70kb/\Delta 70kb}$ mice remains to be elucidated, but it is noteworthy that this noncoding interval appears to influence both CAD risk in humans [1,2] and increased mortality in mice on a high-fat diet through a mechanism that is independent of plasma lipid levels.

Tumors Observed in chr4$^{\Delta 70kb/\Delta 70kb}$ Mice

Seven solid masses that developed spontaneously in chr4$^{\Delta 70kb/\Delta 70kb}$ mice were resected following euthanasia and further examined histopathologically (Charles River Research Animal Diagnostic Services). Six of seven masses were confirmed to be tumors, of which five were different types of sarcomas (Suppl. Table 3). Of note, sarcomas occur spontaneously at increased frequency in *Cdkn2a*$^{INK4a-/-}$ mice [16], possibly indicating a phenotype that is mediated by transcriptional down-regulation of *Cdkn2a* in chr4$^{\Delta 70kb/\Delta 70kb}$ mice. In addition, remarkable hepatosplenomegaly was evident in several chr4$^{\Delta 70kb/\Delta 70kb}$ and chr4$^{+/\Delta 70kb}$ mice. Histological analysis of two chr4$^{\Delta 70kb/\Delta 70kb}$ cases revealed different degrees of extramedullary hematopoiesis, which is again consistent with a phenotype observed in *Cdkn2a*$^{-/-}$ mice [16]. These observations are also interesting in light of the increased susceptibility for other types of cancer associated with the 9p21 locus, e.g. glioma [32,33].

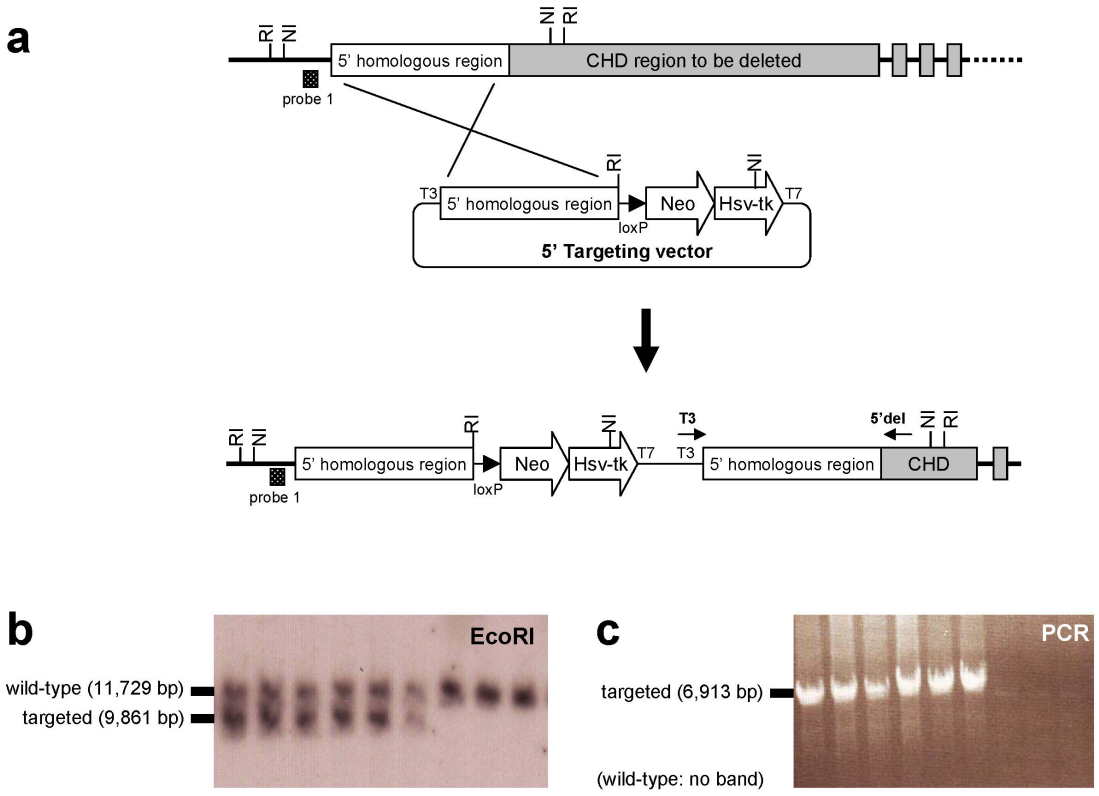
Scanning the Noncoding Risk Interval for Conserved Enhancers

The down-regulation of *Cdkn2a* and *Cdkn2b* in heart and other tissues of chr4$^{\Delta 70kb/\Delta 70kb}$ mice (Fig. 2) in conjunction with the marked differences in expression between the wild-type and chr4$^{\Delta 70kb}$ alleles in heterozygous mice (Fig. 3) provides direct evidence for a *cis*-regulatory function of the noncoding 70kb interval. To further examine if this regulatory function can be attributed to a smaller single enhancer element, we performed comparative genomic analysis of the human 58kb CAD risk interval at chr9p21 and a series of transgenic mouse experiments. To identify enhancer candidate elements, we assessed multi-vertebrate genome alignment data at the UCSC browser [34]. Taking conservation depth, as well as vertebrate evolutionary conservation scores [35] into account, we selected six highly conserved noncoding sequence elements from the larger noncoding interval for transgenic mouse experiments (Suppl. Fig. 9). These sequences were generally well-conserved within the mammalian clade

and had high conservation scores compared to other regions of the risk interval. Based on previous studies using comparative vertebrate genomics for predicting distant-acting enhancers [36], these six sequences were considered to be the most likely subregions of the larger CAD risk interval to be associated with enhancer activity. We cloned all six candidate regions (see Suppl. Table 4 for primer sequences) and tested them using a previously described transgenic mouse reporter assay [37]. In total, 1,243 pronuclear injections (average: 207 per construct) of single-cell stage mouse embryos were performed for these studies. For each construct, we obtained at least 5 (average: 7) LacZ-stained embryos resulting from independent genomic integration events. Embryos were isolated and stained for LacZ activity at embryonic day (e) 11.5 and reporter gene expression patterns were annotated using established reproducibility criteria for this type of assay [37]. In this annotation scheme, elements are only considered to be an enhancer if LacZ staining is observed in the same anatomical structure in at least three embryos resulting from independent transgenic integration events.

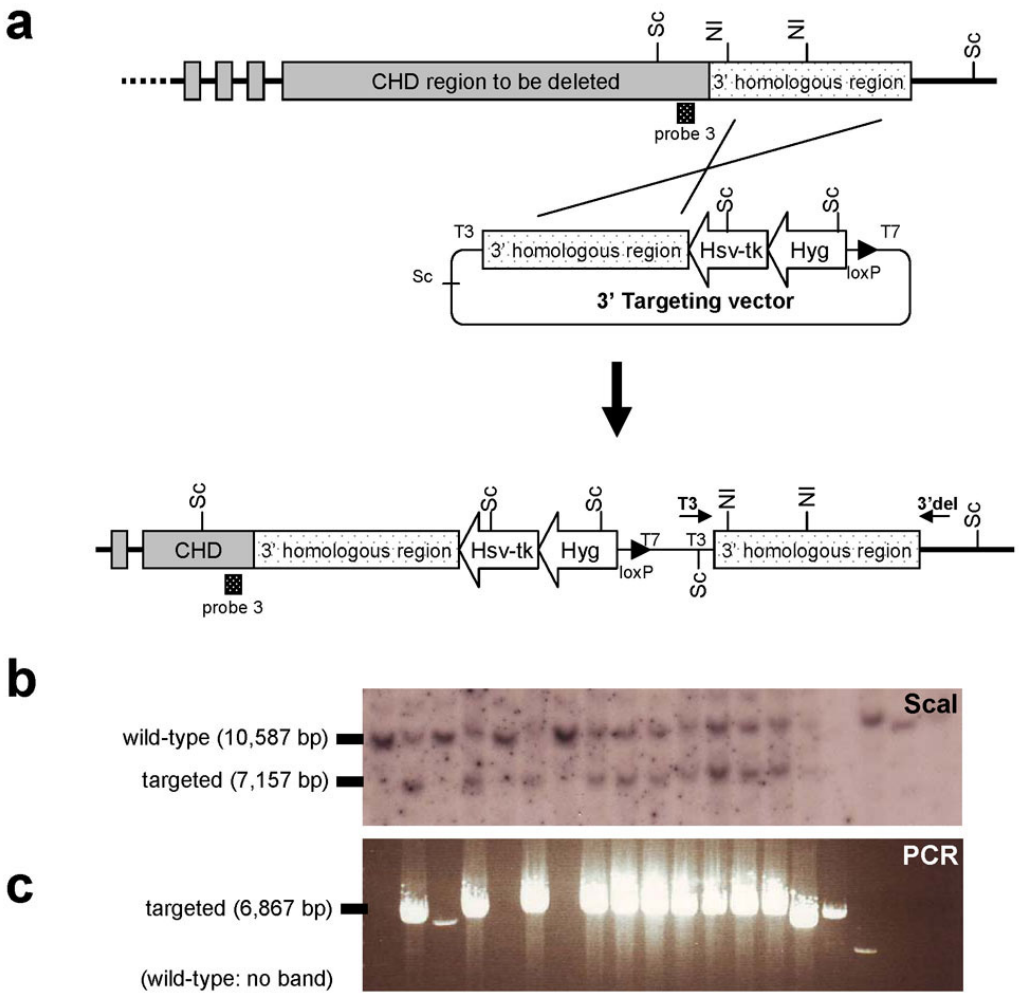
In none of the six cases, a reproducible staining pattern in any embryonic structure (including the heart and vascular system) was observed, indicating that none of the conserved noncoding sequences tested was a reproducible *in vivo* enhancer in this assay at this embryonic stage. In light of the functional evidence from our deletion studies, there are several possibilities to explain this observation. These include: a) one of these sequences could be an enhancer at later time points than e11.5; b) the risk interval contains an enhancer that is not well-conserved in evolution and therefore not easily identified by comparative genomic methods; c) either the enhancer itself or the transcription factors binding to it are not sufficiently conserved between human and mouse to detect *in vivo* activity in this mouse assay using the human candidate sequences; d) the CAD risk interval does not contain any single small subregion (tested elements were 2.1kb-3.2kb in size) that acts a classical enhancer detectable in this assay. This would, e.g., be the case if combinatorial binding of transcription factors to different subregions of the interval was required or if the *in vivo* regulatory effects observed in our deletion studies were primarily due to spacing effects or deletion of non-enhancer types of regulatory element. It is expected that additional experimental studies will be required to pinpoint the exact location and *in vivo* function of such smaller functional elements in the risk interval.

Supplementary Figures



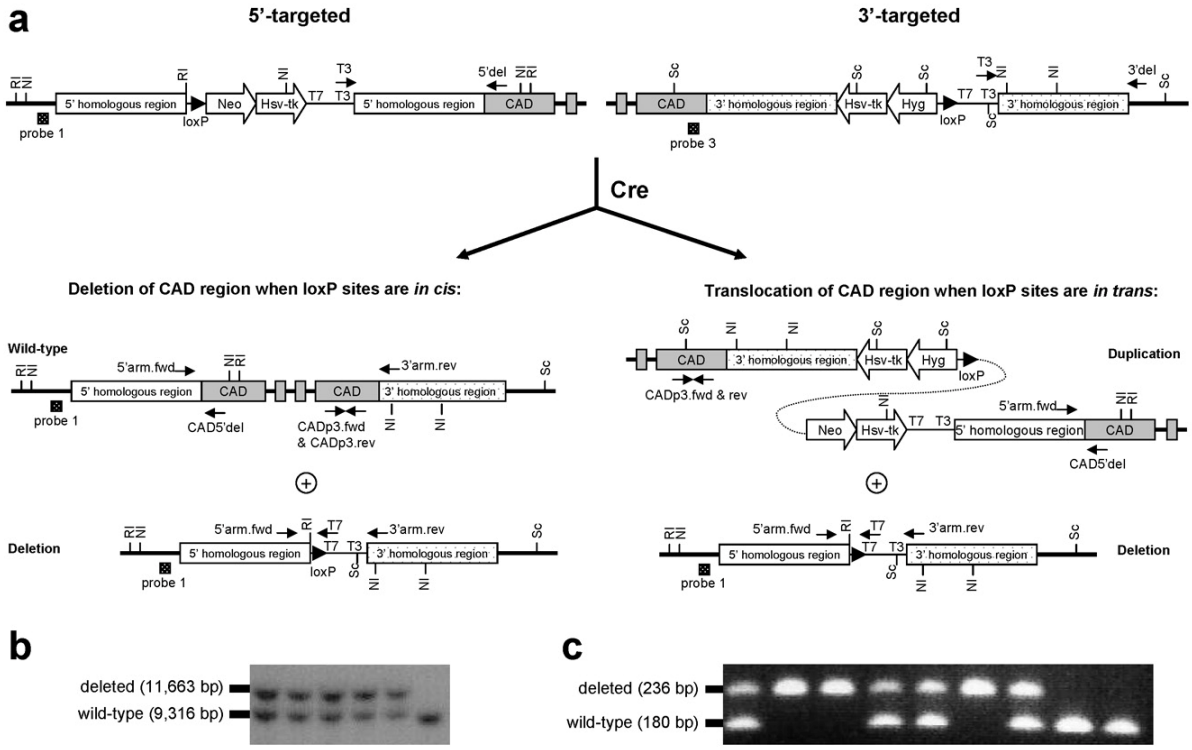
Suppl. Figure 1 – Targeting the 5'-end of the 70kb risk interval

a) Schematic strategy for introduction of a LoxP site near the 5'-end of the region of interest. b) Validation of successful recombination by Southern hybridization. c) Validation of successful recombination by PCR. See Methods for details.



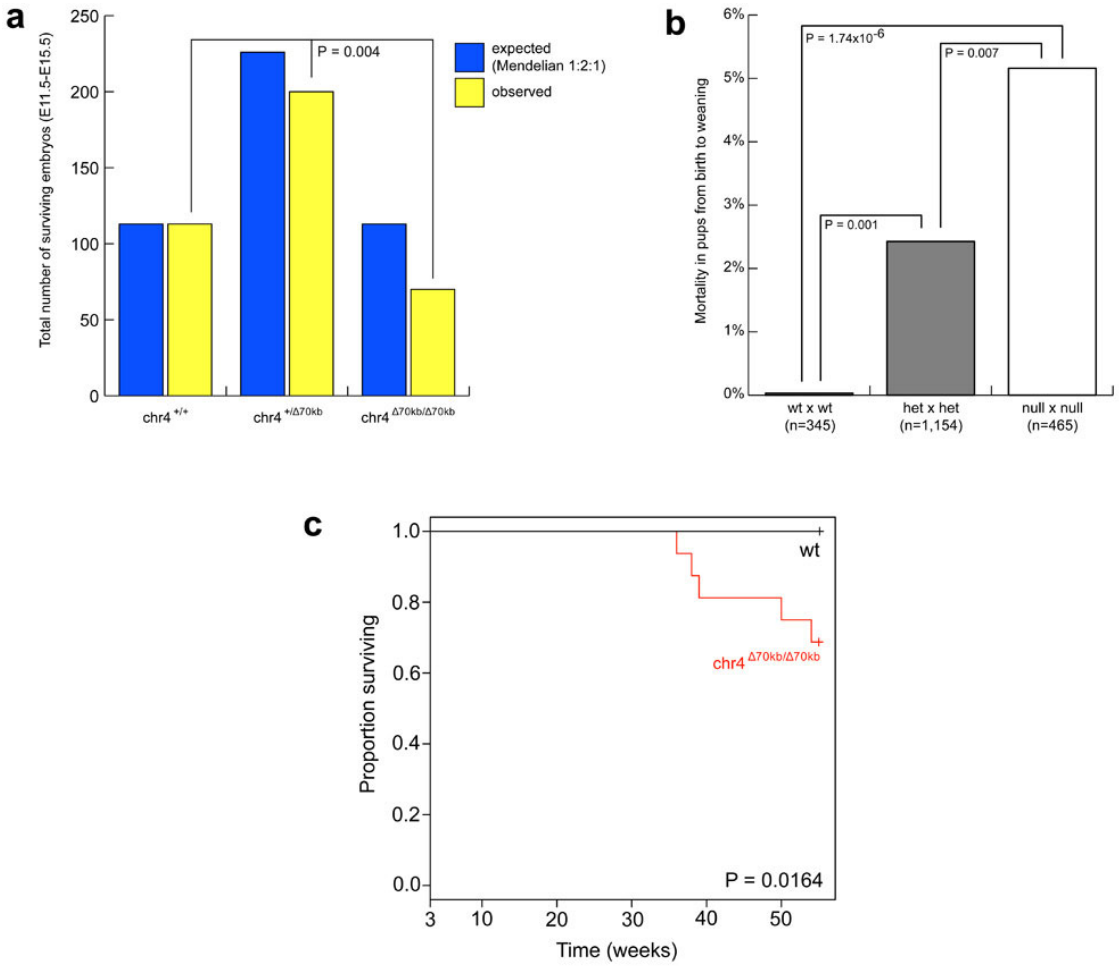
Suppl. Figure 2 – Targeting the 3'-end of the 70kb risk interval

a) Schematic strategy for introduction of a LoxP site near the 3'-end of the region of interest. b) Validation of successful recombination by Southern hybridization. c) Validation of successful recombination by PCR. See Methods for details.



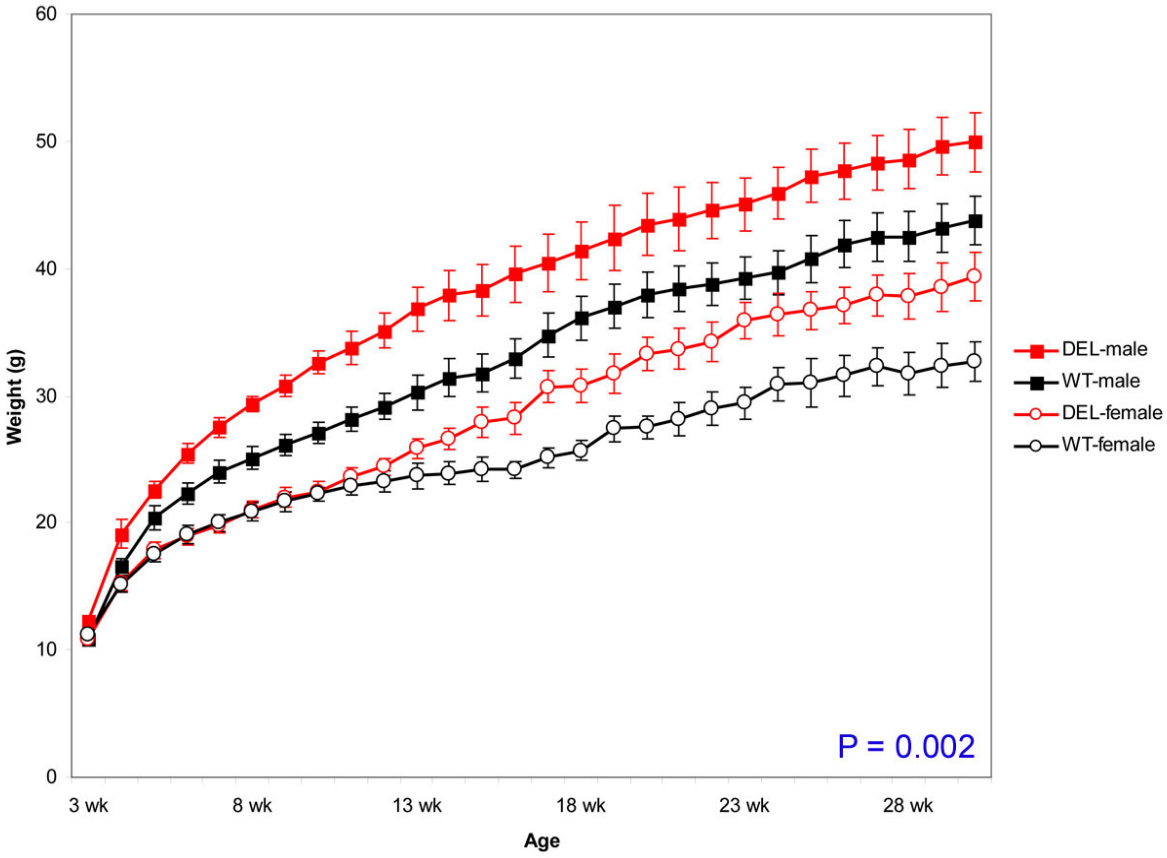
Suppl. Figure 3 – Deletion of 70kb-interval by Cre-mediated recombination after double-targeting

a) Schematic strategy for recombination of double-targeted loci, resulting in deletion of the 70kb region of interest. b) Validation of successful deletion by Southern hybridization. c) PCR genotyping results of wild-type, heterozygous and homozygous chr4$^{\Delta 70kb}$ mice from crosses of chimera-derived heterozygous founders.



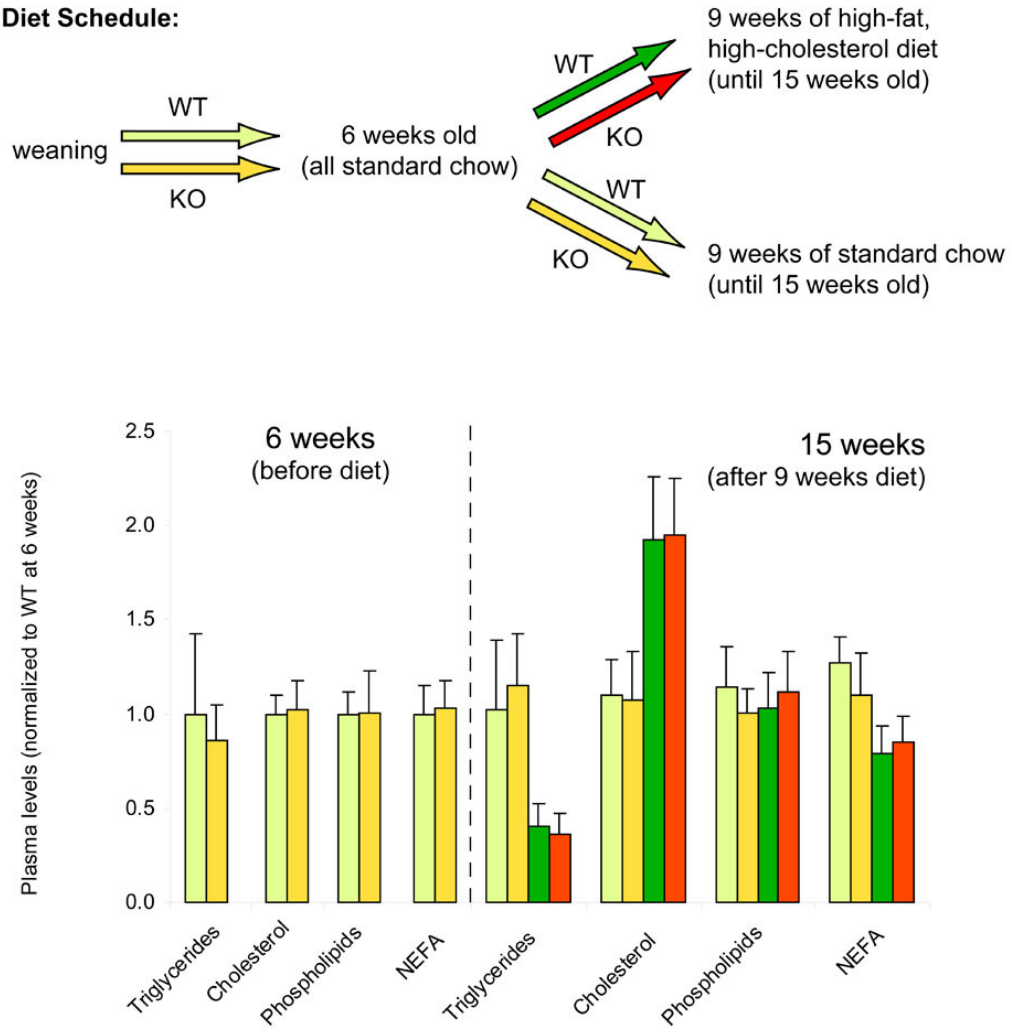
Suppl. Figure 4 – Decreased embryonic, postnatal and adult survival

a) chr4$^{+/\Delta 70kb}$ and chr4$^{\Delta 70kb/\Delta 70kb}$ genotypes are underrepresented among morphologically normal, non-aborted embryos collected between embryonic days E11.5 and E15.5, indicating reduced survival compared to wild-type embryos. b) Pups from heterozygous and homozygous chr4$^{\Delta 70kb}$ crosses are more likely to die between birth and weaning. c) Chr4$^{\Delta 70kb/\Delta 70kb}$ mice are less likely to survive up to 55 weeks of age than wild-type controls. *P*-values: G-test (a), two-tailed Fisher's Exact test (b), Kaplan-Meier test (c).



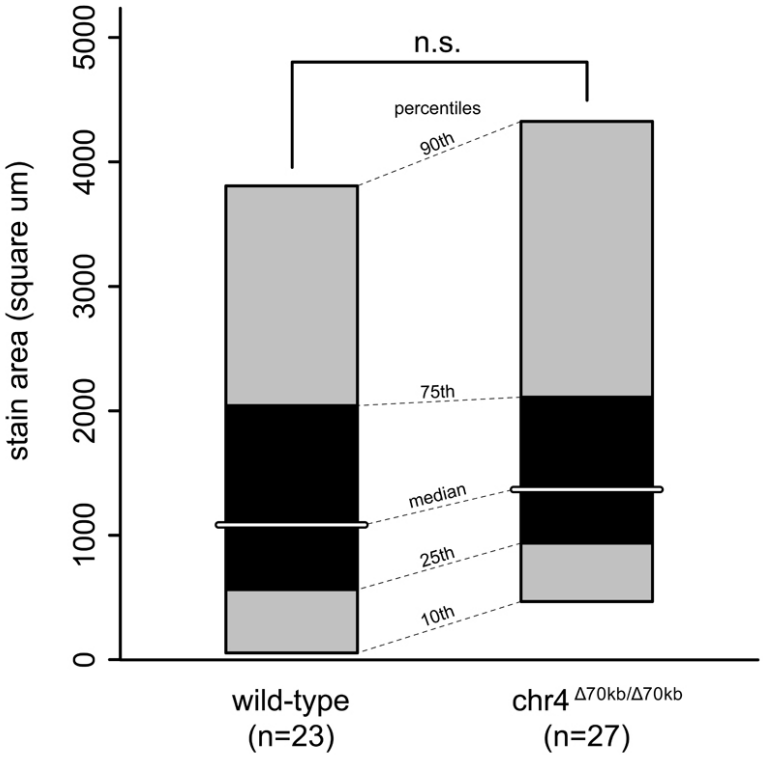
Suppl. Figure 5 – Chr4$^{\Delta 70kb/\Delta 70kb}$ mice have increased body weight

Growth curve on standard chow (*ad libitum*) up to 30 weeks of age. Plotted values are means ± SEM for 8 animals per gender and genotype. P=0.002 (t-test, two-tailed, paired, across males and females at 30 weeks of age). No significant difference from wild-types was observed for heterozygotes (not shown).



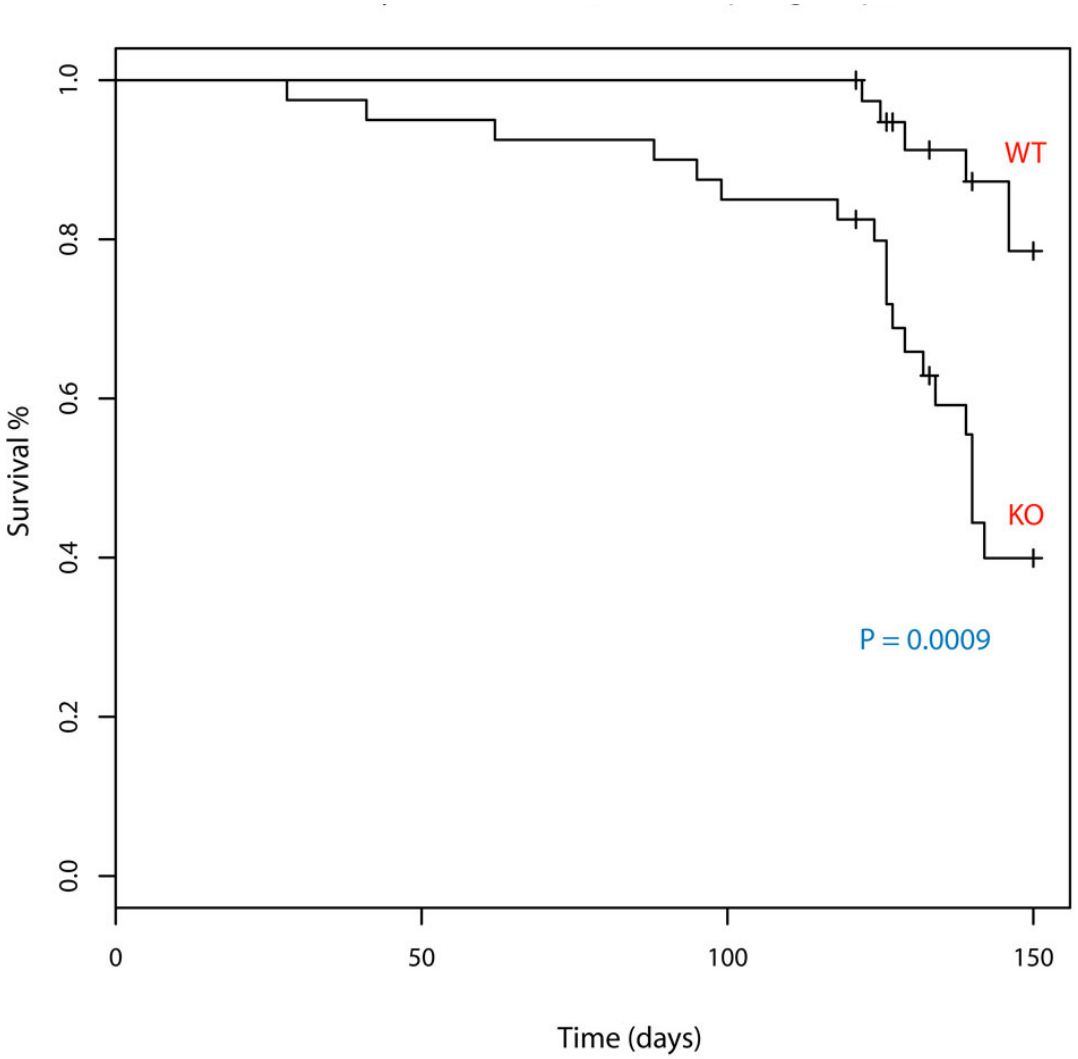
Suppl. Figure 6 – Chr4$^{\Delta 70kb/\Delta 70kb}$ mice have normal plasma lipid levels

Deletion mice and wild-type controls were placed on standard chow up to 6 weeks of age. Control groups were kept on standard chow for an additional 9 weeks, the test groups were placed on a high-fat, high-cholesterol diet for 9 weeks. No significant differences between wild-type and deletion mice was observed in base lipid levels or in the response to the noxious diet. NEFA = non-esterified free fatty acids.



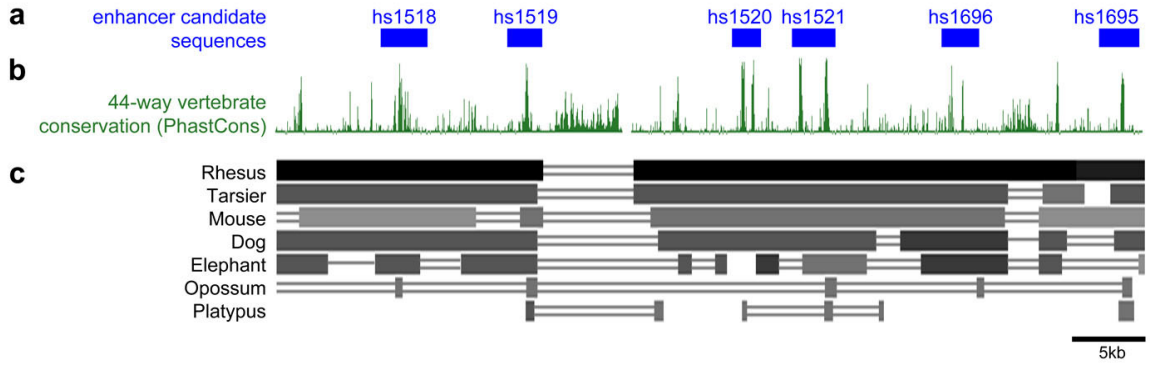
Suppl. Figure 7 – Chr4$^{\Delta 70kb/\Delta 70kb}$ mice show no increase in atherogenic lesions

Deletion mice and isogenic controls in a 129Sv background were placed on a high-fat, high-cholesterol diet for ~20 weeks and assayed for atherogenic lesions. Extent of atherogenic lesions varied over a wide range and no significant increase in atherogenesis was observed in deletion mice (P=0.28, Mann-Whitney U test, non-directional).



Suppl. Figure 8 – Increased mortality on high-fat, high-cholesterol diet

Survival curve of wild-type and chr4$^{\Delta 70kb/\Delta 70kb}$ mice on high-fat, high-cholesterol diet fed *ad libitum* for 150 days. See Methods for details about diet and euthanasia criteria. P-value: Kaplan-Meier test for survival up to 150 days on high-fat diet.



Suppl. Figure 9 – Conserved noncoding sequences within the CAD risk interval

a) Six conserved noncoding sequences (blue boxes) located within the human CAD risk interval were identified based on 44-way vertebrate genome sequence alignments at the UCSC genome browser [34]. In a transgenic mouse enhancer assay [38], no reproducible enhancer activity in day 11.5 mouse embryos was observed in the heart or any other structure. b) 44-way vertebrate conservation across the 58kb noncoding CAD risk interval [35]. c) Conservation in selected vertebrate species.

Supplementary Tables

Primer name	Primer sequence	Product	Note
CHD 5' fwd	ACTTCAGTAGGATCCTGATTTTGAGGTACTTTTTAGACAGTTTAGAAATG	6223 bp	generate homology sequence
CHD 5' rev	ACTTCAGTAGAATTCAGGAAGAACCACGACTCCACATACTTG		for 5' knock-in
CHD 3' fwd	ACTTCAGTAGCGGCCGCTTCAGGGCCAGAGCTTCATAATGAAATAGT	6284 bp	generate homology sequence
CHD 3' rev	ACTTCAGTAAGATCTGAGCCAAGGTACTGCTATTGTTGTTTGC		for 3' knock-in
T3	CGCAATTAACCCTCACTAAAGGGAAC	6913 bp	PCR screening for 5' knock-in
CHD5'del	TGCACGGTCAATGGTTTCTCAATGCC		wild-type no band
T3	same as above	6867 bp	PCR screening for 3' knock-in
CHD3'del	GCAGAGAAGCAAAGCTGGTTTTCACA		wild-type no band
CHDp1.fwd	TGGAGTCTTCAGAAACTTGTCACATACTTC	361 bp	probe 1
CHDp1.rev	CACATCCCGATCCAAATATAATTCTAGCCT		for southern
CHDp3.fwd	AAGGTATCCTAAATACTGTCTTCTTGCAG	180 bp	probe 3
CHDp3.rev	CGAGTCAATTTTCTTCATGTTTATCCTCCA		for southern
T7	CGTAATACGACTCACTATAGGGCG	236 bp	PCR screening deletion event
CHD5'arm.fwd	TATGAAAGCACACTTGTGGGCGTGT		wt: no band; deletion: 236 bp
CHD5'arm.fwd	same as above	3136 bp	PCR confirming deletion event
CHD3'arm.rev	TGTACCAGAAAGGACAATGAACTCCTTGAT		wt: no band; deletion: 3136 bp
T7	same as above	wt-180 bp	for PCR genotyping mice
CHD5'arm.fwd	same as above	del-236 bp	mix primers at 1:1:1:1 ratio
CHDp3.fwd	same as above		
CHDp3.rev	same as above		

Suppl. Table 1 - Primer sequences used for generation of targeting constructs and genotyping

Primer name	Primer sequence	Product
Cdkn2b-F	AGATCCCAACGCCCTGAAC	110 bp
Cdkn2b-R	CGCAGTTGGGTTCTGCTC	
p16-INK4aF	CCCAACGCCCCGAACT	79 bp
p16-INK4aR	GCAGAAGAGCTGCTACGTGAA	
Mtap E2f	TGGTGGAACAGGCTTGGATGATCC	164 bp
Mtap E4r	AAGGCATGATGGTGTGTTGTCTGCC	
Dmrta1 E1f	TGCCTCTAGACACCCTGGGAGC	146 bp
Dmrta1 E2r	GGTGATGAGTGTTGGAGACTGGTTCTTC	
CDKN2B-rtSNP-F	GTTTTCCCAGTCACGACGTTGTAAAGAGCAACTCAAATGTAGGAAA	1019 bp
CDKN2B-rtSNP-R	AGGAAACAGCTATGACCATAGATCCCAACGCCCTGAAC	
CDKN2B-Genomic-F	GTTTTCCCAGTCACGACGTTGTAAAGAGCAACTCAAATGTAGGAAA	470 bp
CDKN2B-Genomic-R	AGGAAACAGCTATGACCATGGCCCTCTACCTTTCAGGAC	

Suppl. Table 2 – Primer sequences used for quantitative RT-PCR and allele-specific expression profiling

Genotype	Age (weeks)	Sex	Site of tumor/mass	Histopathological diagnosis
chr4$^{\Delta 70kb/\Delta 70kb}$	51	M	skin	hemangiosarcoma
chr4$^{\Delta 70kb/\Delta 70kb}$	54	F	skin	anaplastic sarcoma with features of hemangiosarcoma (necropsy also revealed an irregular liver, histological evaluation confirmed myeloid leukemia)
chr4$^{\Delta 70kb/\Delta 70kb}$	28	F	leg	osteosarcoma
chr4$^{\Delta 70kb/\Delta 70kb}$	28	F	skin	sarcoma with moderate eosinophilic inflammation
chr4$^{\Delta 70kb/\Delta 70kb}$	55	M	lung	sarcoma with features of hemangiosarcoma
chr4$^{\Delta 70kb/\Delta 70kb}$	61	F	uterus	metritis
chr4$^{\Delta 70kb/\Delta 70kb}$	15	M	abdomen	teratoma

Suppl. Table 3 – Summary of histopathological analysis of tumors incidentally found in chr4$^{\Delta 70kb/\Delta 70kb}$ mice.

Candidate sequence	Coordinates (hg18)	Primer sequence	Fragment size
hs1518	chr9:22068885-22072105	Forward: CTTTTGGGTTTCCCCATTGT Reverse: AACAAGTGAACTGGGGACCA	3221bp
hs1519	chr9:22077552-22079798	Forward: GCCTAGTGGAAATTTCTATTGCTG Reverse: GTCATTGGCTCAATCTAATACCAA	2247bp
hs1520	chr9:22092664-22094753	Forward: CACTCACCTAAAACCCAAAAACA Reverse: CAATGCCTGGCACCTAGAAT	2090bp
hs1521	chr9:22096994-22099800	Forward: AGGATAGTCTGCATTTCATGGT Reverse: CCACTTTAGGTTCCCCACAA	2807bp
hs1695	chr9:22117677-22120448	Forward: GAGACAGGAGGGTCCCAAAT Reverse: AGAGGAATCACACCTCTGGAA	2772bp
hs1696	chr9:22106987-22109574	Forward: CAAATGGAAGCTGGGAGTGT Reverse: GAAGGATGGTCATTGTTCCA	2588bp

Suppl. Table 4 – Primer sequences and coordinates of candidate enhancer sequences that were tested in transgenic mouse embryos.

References (Supplementary Material)

[32] Shete, S. *et al.*, Genome-wide association study identifies five susceptibility loci for glioma. *Nat Genet* 41 (8), 899-904 (2009).

[33] Wrensch, M. *et al.*, Variants in the CDKN2B and RTEL1 regions are associated with high-grade glioma susceptibility. *Nat Genet* 41 (8), 905-908 (2009).

[34] Kuhn, R.M. *et al.*, The UCSC Genome Browser Database: update 2009. *Nucleic Acids Res* 37 (Database issue), D755-761 (2009).

[35] Siepel, A. *et al.*, Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* 15 (8), 1034-1050 (2005).

[36] Visel, A. *et al.*, Ultraconservation identifies a small subset of extremely constrained developmental enhancers. *Nat Genet* 40 (2), 158-160 (2008).

[37] Pennacchio, L.A. *et al.*, In vivo enhancer analysis of human conserved non-coding sequences. *Nature* 444 (7118), 499-502 (2006).

[38] Visel, A., Minovitsky, S., Dubchak, I., & Pennacchio, L.A., VISTA Enhancer Browser--a database of tissue-specific human enhancers. *Nucleic Acids Res* 35 (Database issue), D88-92 (2007).