



**Figure S2** Genetic variation in zebrafish strains detected in low coverage WGS data of pooled DNA from 20 fish. **(a)** Pairwise comparison of SNP genotypes between parental lines showing the percentages of SNPs that are polymorphic and thus could be used to predict the parental origin for mapping based on homozygosity-by-descent. SNPs were classified as polymorphic if only the reference genome allele was observed in one line, while at least one alternate allele was observed in the second line. **(b)** Graph showing the number of polymorphic SNPs (in millions) identified in each parental line, as well as the number of lines with which these SNPs are shared. For (a) and (b), only sites with sequence coverage in all lines were considered (5.2 million sites of the 7.6 million total SNP sites). **(c, d)** Venn diagrams showing the SNPs that were classified as mapping strain SNPs. This includes 0.74 million SNPs found in TLF but not in either of the Tü lines, and 1.2 million SNPs found in either of the WIK lines but not in either of the Tü lines. For (c) the two Tü lines are shown separately, while in (d) the data from these two lines were combined. Interestingly, due to high levels of intra-strain variation, there is a high number of SNPs that are not shared between the two Tü lines (c), and a similar number of SNPs that are not shared between the two WIK lines (d).