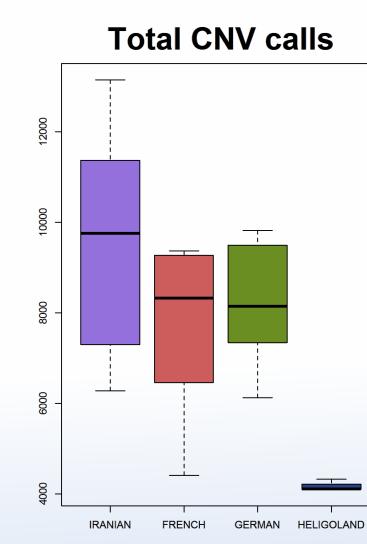
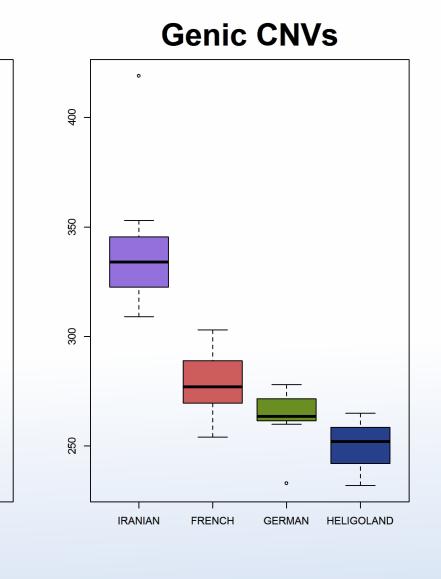
Data set 27 house mice were sequenced by using Illumina HiSeq technology. Three groups of 8 individuals derived from wild caught mice in three natural populations: Cologne/ Bonn (German), Massif Central (French) and Ahvaz (Iran).

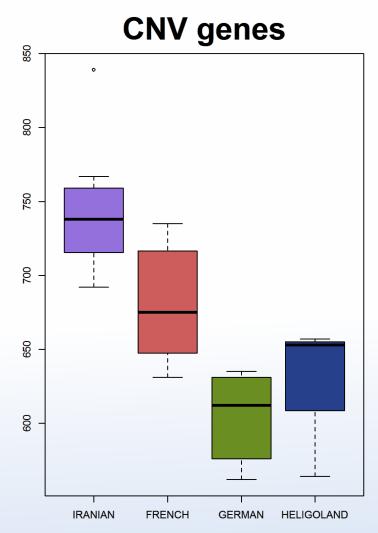
Three additional individuals were caught on Heligoland island in the North Sea.

CNV discovery

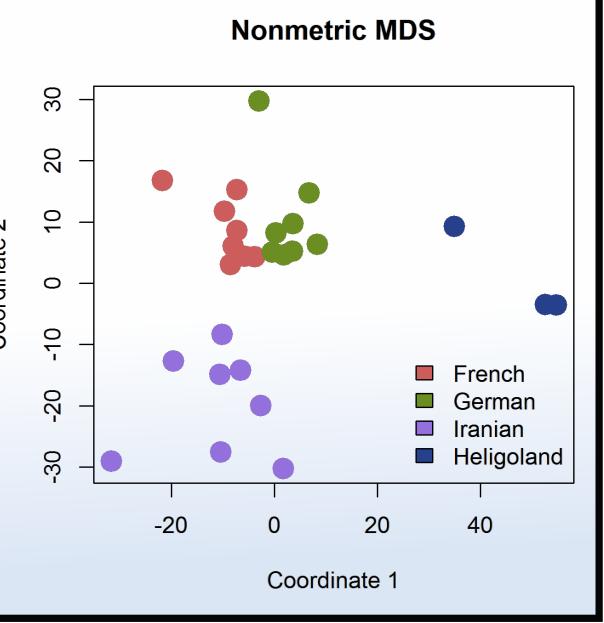
We used CNVnator software to predict CNV regions¹. CNV calls were intersected with RefSeq gene coordinates to obtain only those transcription units entirely confined within predicted CNVs (1,925 genes).





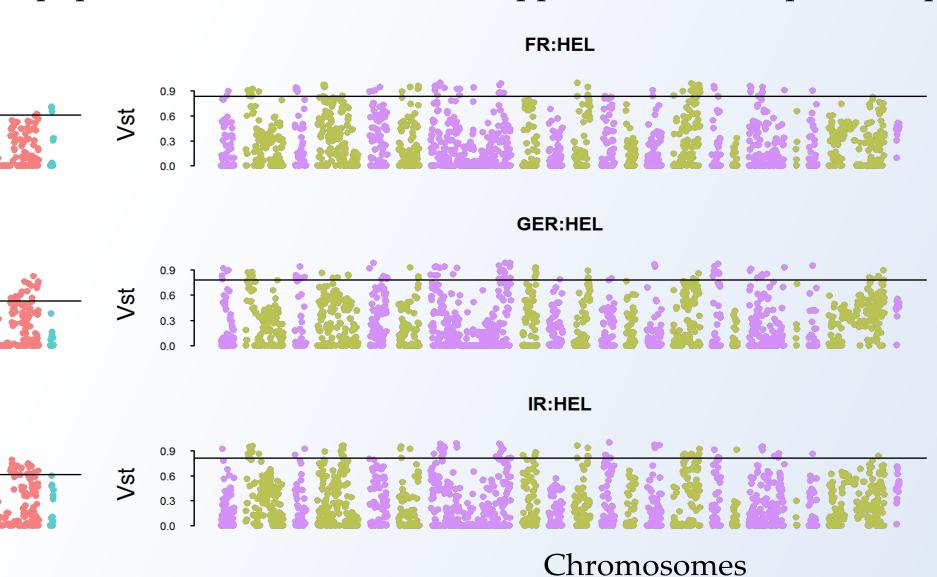


Multidimensional scaling ordination of mouse populations, based on dissimilarity matrix generated by calculating the Euclidian distance between each of the possible pairwise sample comparisons of 27 individuals, across all 1,925 CNV genes.

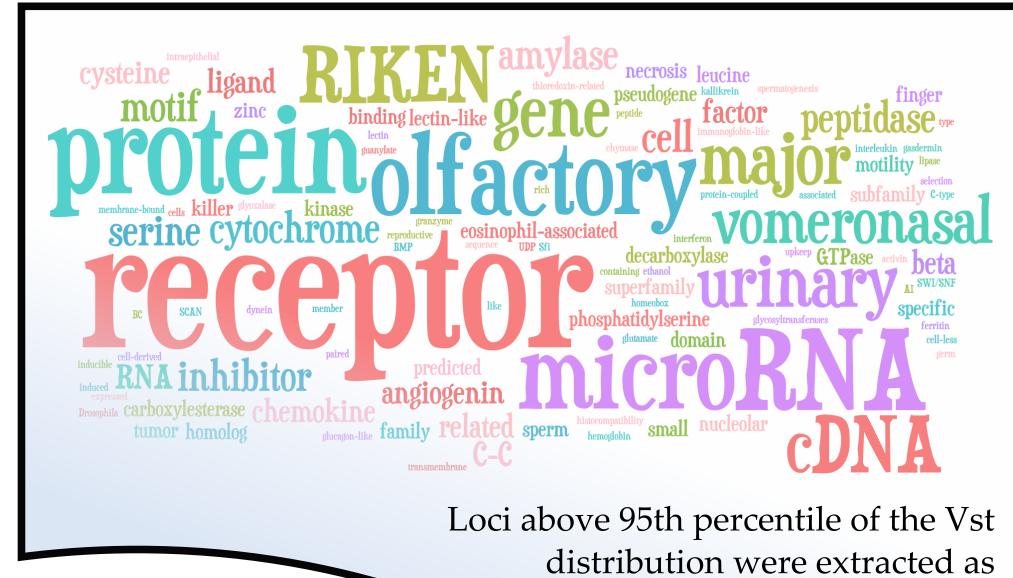


Population differentiation

To identify copy number variable genes showing high differentiation between populations, Vst statistics was applied². For each pairwise population comparison,



Vst was calculated for every CNV gene as Vst=(Vt-Vs)/Vt where Vt is the variance between the two populations and Vs is average of the variance within each population, weighted for its sample size.



the most differentiated within comparison. The majority belongs to enviromentally responsive genes.

Population genetics of

CNV genes
in natural mouse populations

Max Planck Institute for Evolutionary Biology

Chromosomes

Plön, Germany pezer@evolbio.mpg.de

MAX-PLANCK-GESELLSCHAFT

Željka Pezer Sakač Bettina Harr Meike Teschke Hiba Babiker Diethard Tautz

Genes identified

to have significant

difference in copy

number between

two populations

(p<0.05, Tukey's

calculate mean

frequencies of

CN ratios.

ratios for

hoc test.

test) were used to

Histograms show

log2 transformed

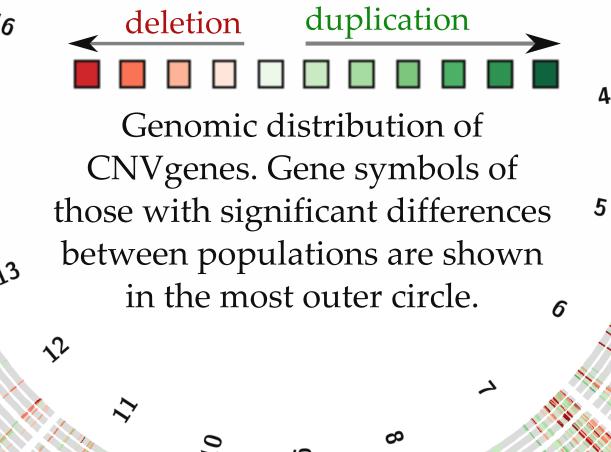
indicated number

of genes (n) that

passed the post

Introduction All multicellular species examined

thus far display differences between individuals in their gene content. These genic copy number variations (CNVs) are likely to be functionally important and have been shown in some cases to underlie adaptive phenotypic variation. CNVs can affect fitness of an organism through various molecular mechanisms such as gene dosage or position effect, and can therefore drive gene and genome evolution. Here, we investigated gene copy number polymorphism in wild mice and its effect on population structure.



log2 transformed values of the population average copy number for each CNV gene were plotted between populations.

The most

differentiated

distribution.

genes are shown

in red as outliers

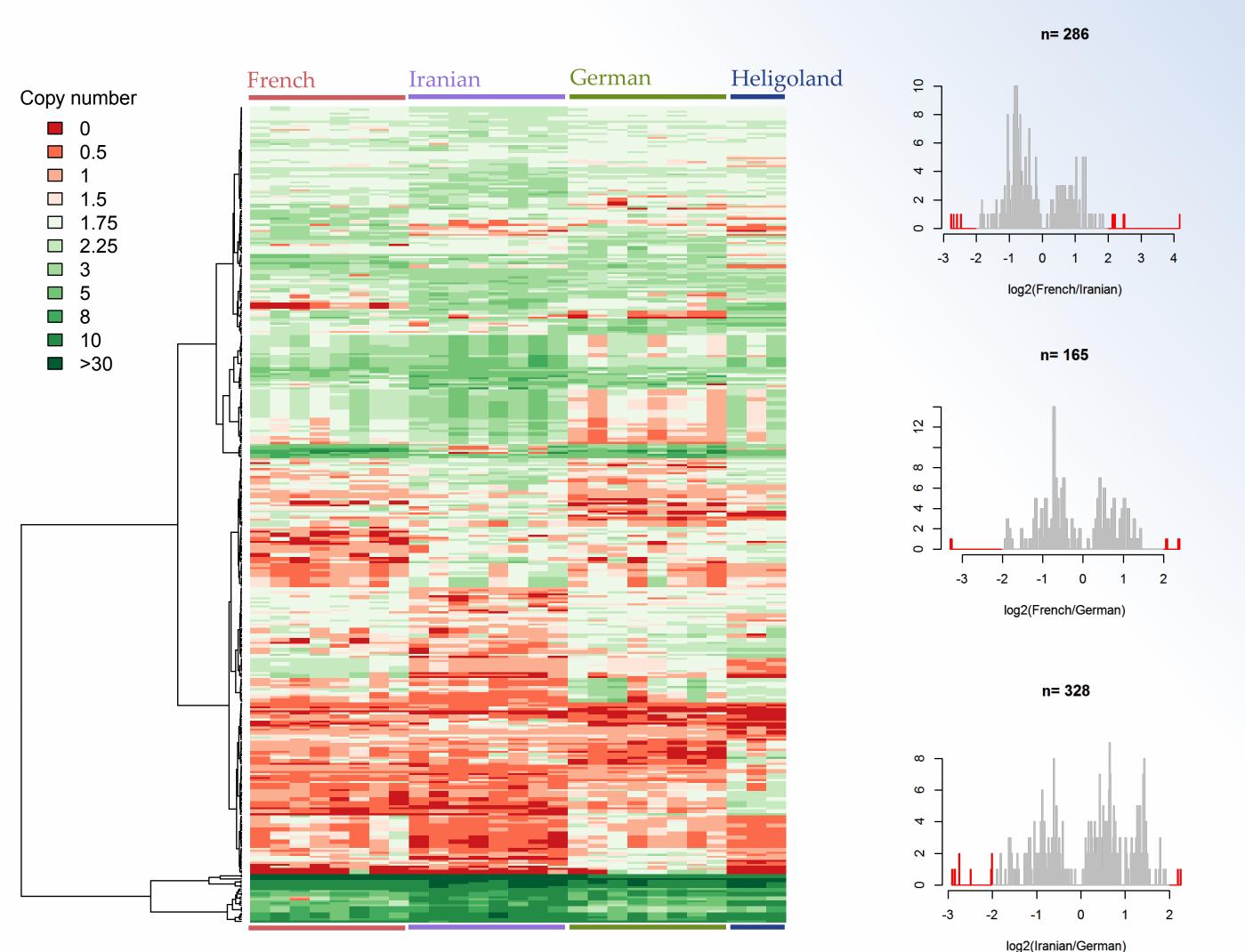
from the resulting

2. Redon R et al., Nature 444:444-454 (2006)

CNV genes with significant population differences

We applied Anova analysis on the whole CNV gene set between French, German and Iranian population. Heligoland population was excluded from the analysis because its sample size was too small and unequal to the other three populations.

Heatmap of genes showing significant differences in copy number between populations. Copy number states of 404 genes (in rows) that were significantly different in at least one pairwise comparison (p<0.05, Anova+FDR). Each column represents one individual.



Pairwise population comparison of mean copy number (log2) across loci

Genes with significant difference in population mean copy number Gm10512 Smok2a Smok2b Gm14850 French Heligoland Iranian

1. Abyzov A *et al.*, Genome Res 21:974–984 (2011)