

Copy number variants in discontinuous landscapes of heterozygosity in the mouse genome

Hailie Pavanel¹, Nicholas Boehler¹, Bin Luo², R. J. Kulperger², C. B. Dean³ and Kathleen A Hill¹, khill22@uwo.ca

¹Department of Biology, University of Western Ontario, London ON, Canada; ²Department of Statistical and Actuarial Science, University of Western Ontario, London ON, Canada;

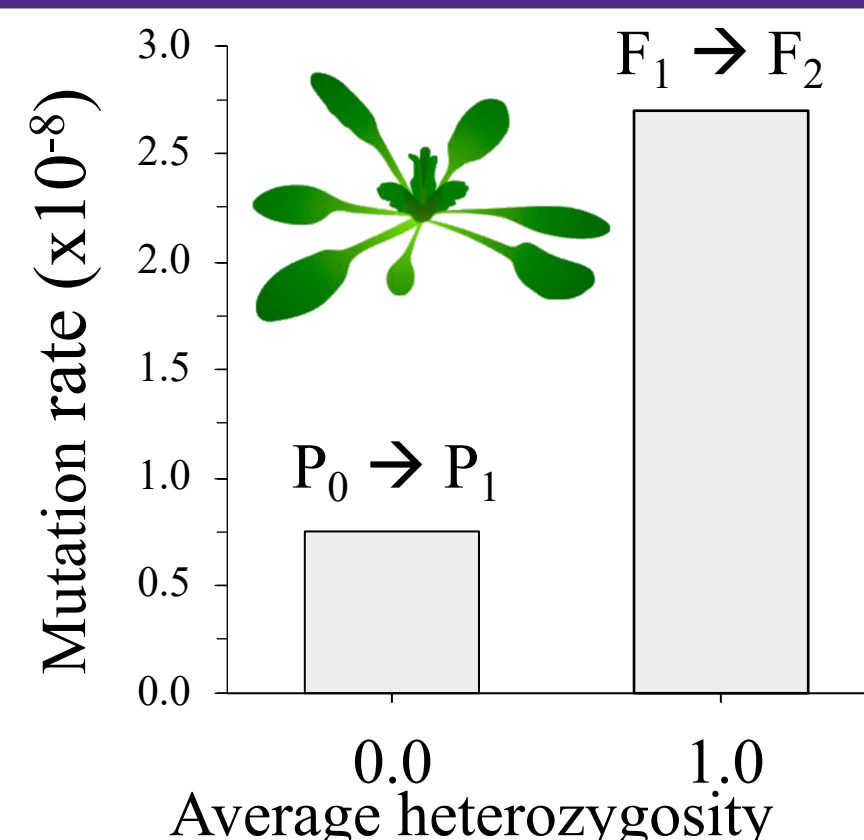
³Department of Statistics and Actuarial Science, University of Waterloo, Waterloo ON, Canada

Higher mutation rates are associated with localized regions of higher heterozygosity

The rate of mutation is not constant along the chromosome sequence and mechanisms are not fully understood.

Increased mutagenesis has been associated with heterozygosity in *Arabidopsis* but this phenomenon has yet to be fully explored in mammals.

Yang *et al* 2015 *Nature* 523:463

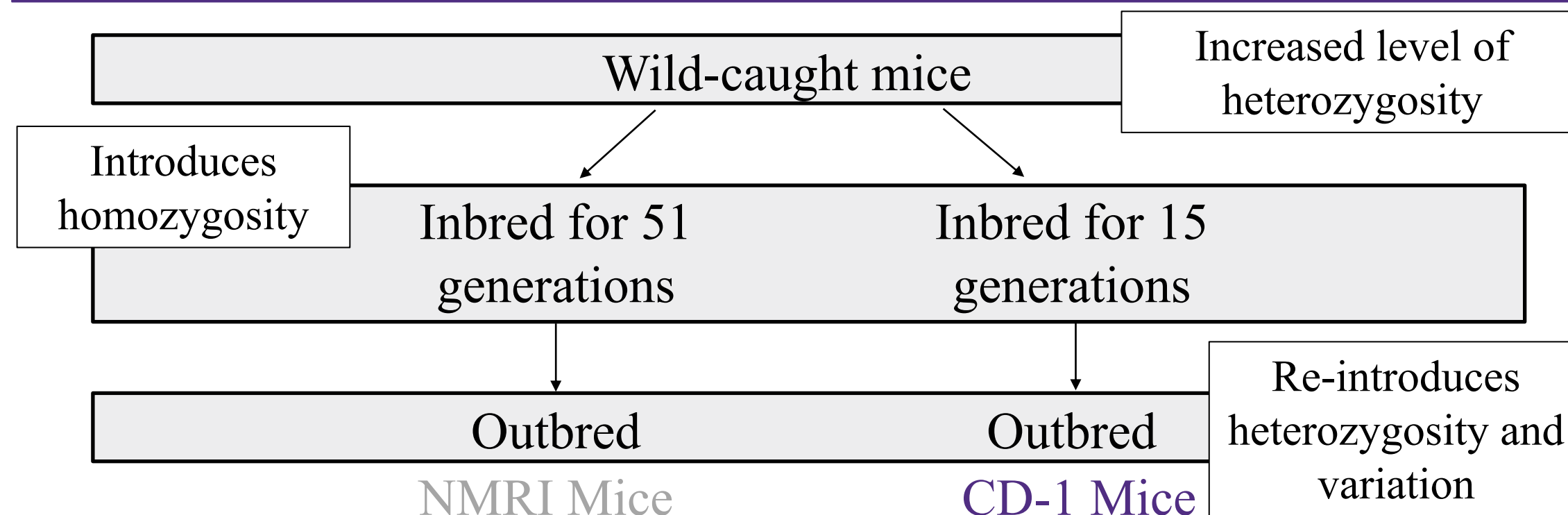


Inbred mice are often used as models for mutation research but their genetic background precludes analysis of the impact of heterozygosity.

Adapted from Yang *et al* 2015 *Nature* 523:463

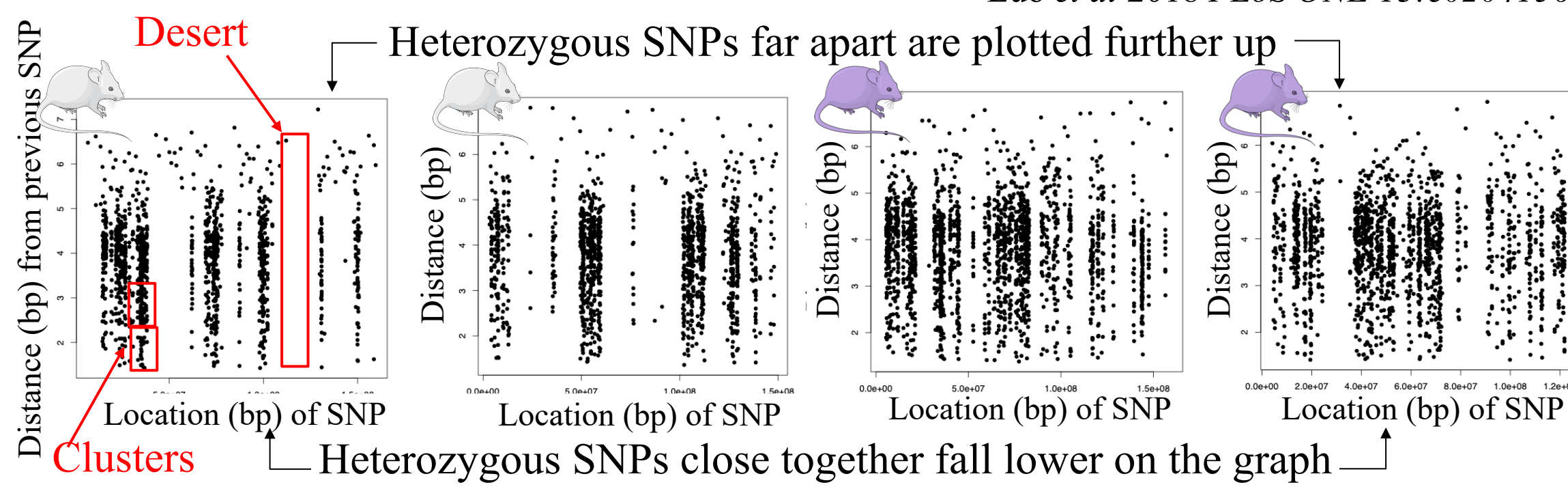
Two commercially-available stocks of outbred mice, NMRI (Naval Medical Research Institute) and CD-1 (Caesarian Derived-1) provide the opportunity to examine mutagenesis in a more heterozygous landscape.

Background of outbred mice breeds a discontinuous landscape of heterozygosity



We use **rainfall plots** to display the spatial distribution of heterozygous single nucleotide polymorphisms (SNPs) and to identify clusters of heterozygous SNPs and deserts, where there are no heterozygous SNPs.

Luo *et al* 2018 *PLoS ONE* 13:e0204156



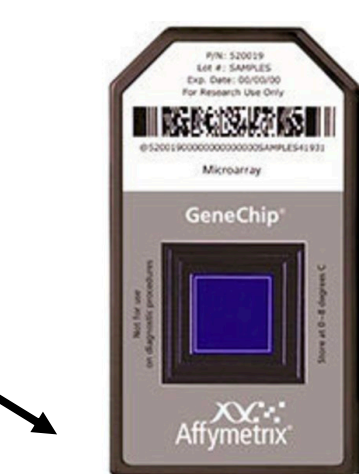
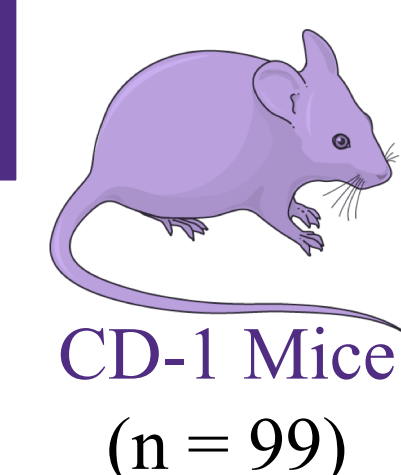
Research question: does heterozygosity attract mutations?

Statistical tools to assess the spatial association between two mutational events

We use spatial statistic tools to analyze the genomic context to associate heterozygous single nucleotide polymorphisms (SNPs) and copy number variants (CNVs).



Detection of CNVs and heterozygosity in CD-1 and NMRI mice and assessing spatial association using J statistic



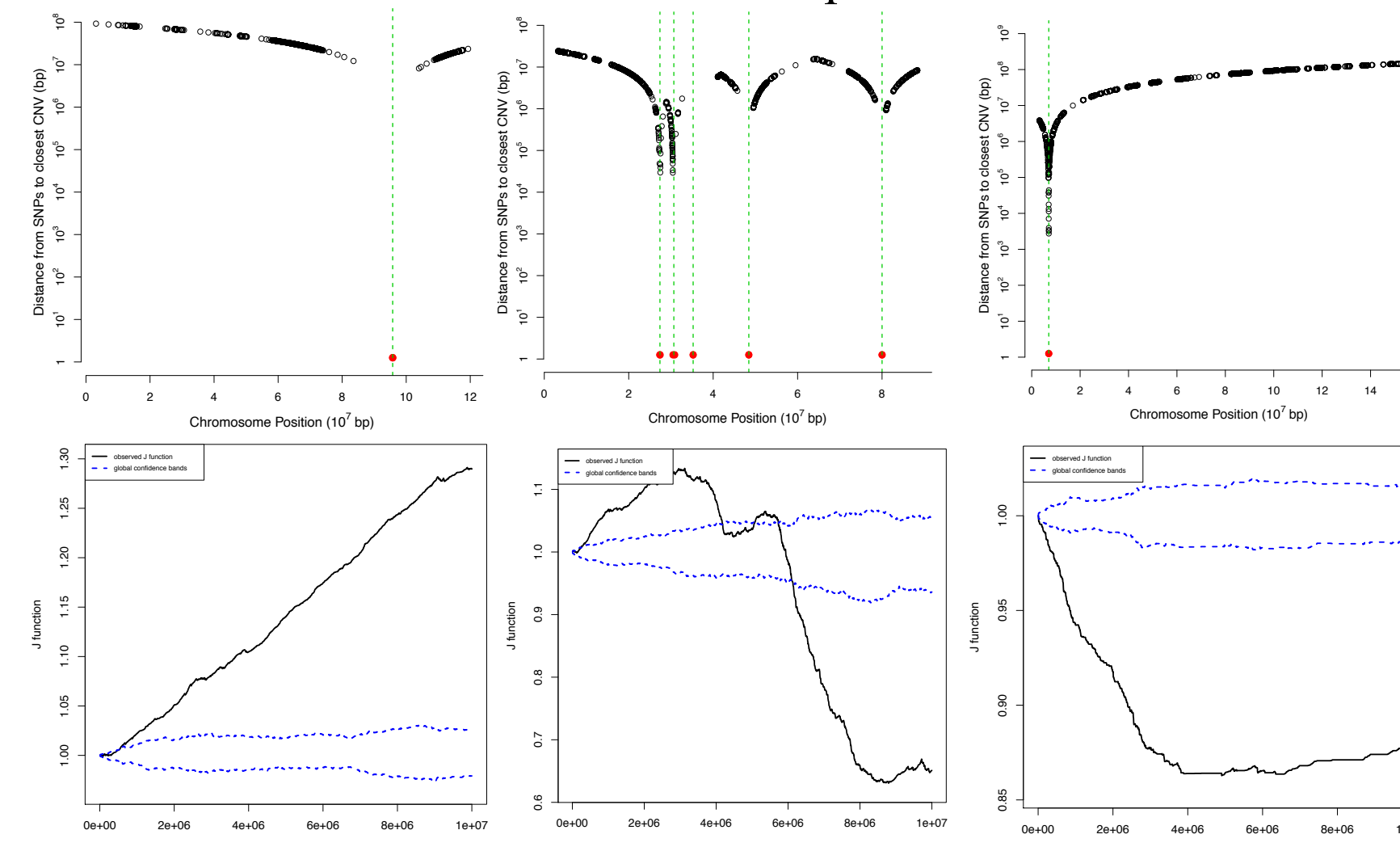
Detection of base substitutions (SNPs) and copy number variants (CNVs)

Three possible outcomes from J statistic:

Distal: CNV occurs far away from heterozygous SNPs

Both: CNV occurs first close to or far from heterozygous SNPs and then the landscape varies

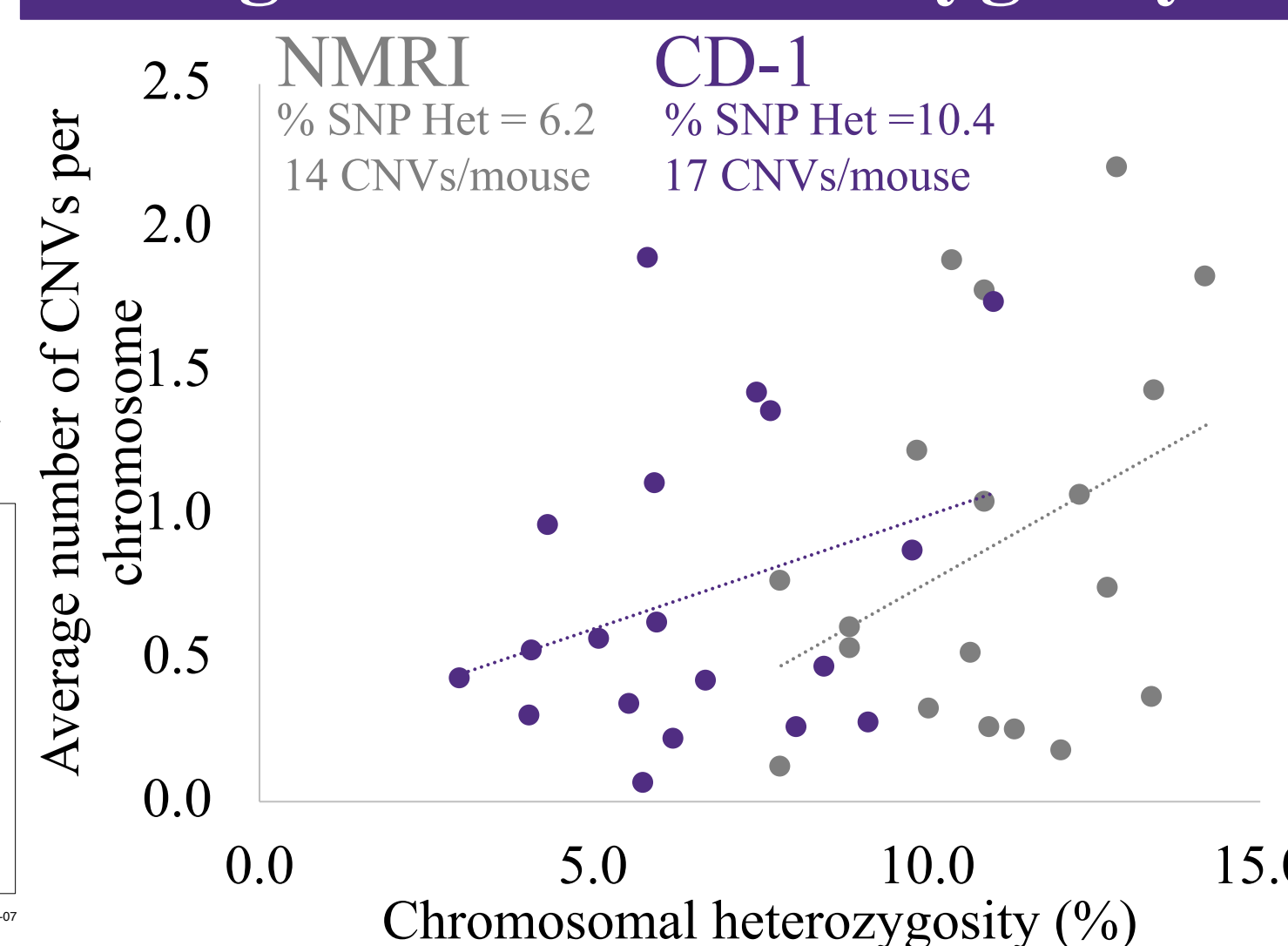
Proximal: CNV occurs close to heterozygous SNPs



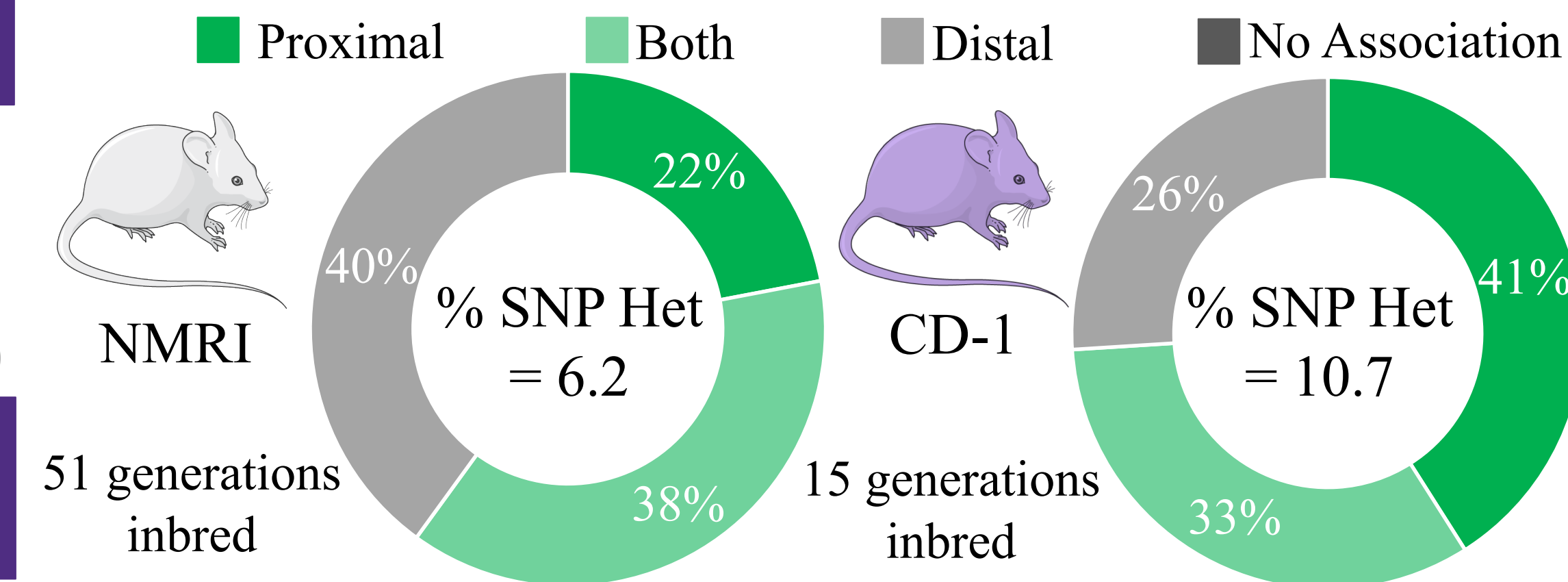
We use **rainbow plots** to visualize spatial association between heterozygous SNPs and CNVs

We use the **J statistic** to assess the significance of the association after running 1000 simulations

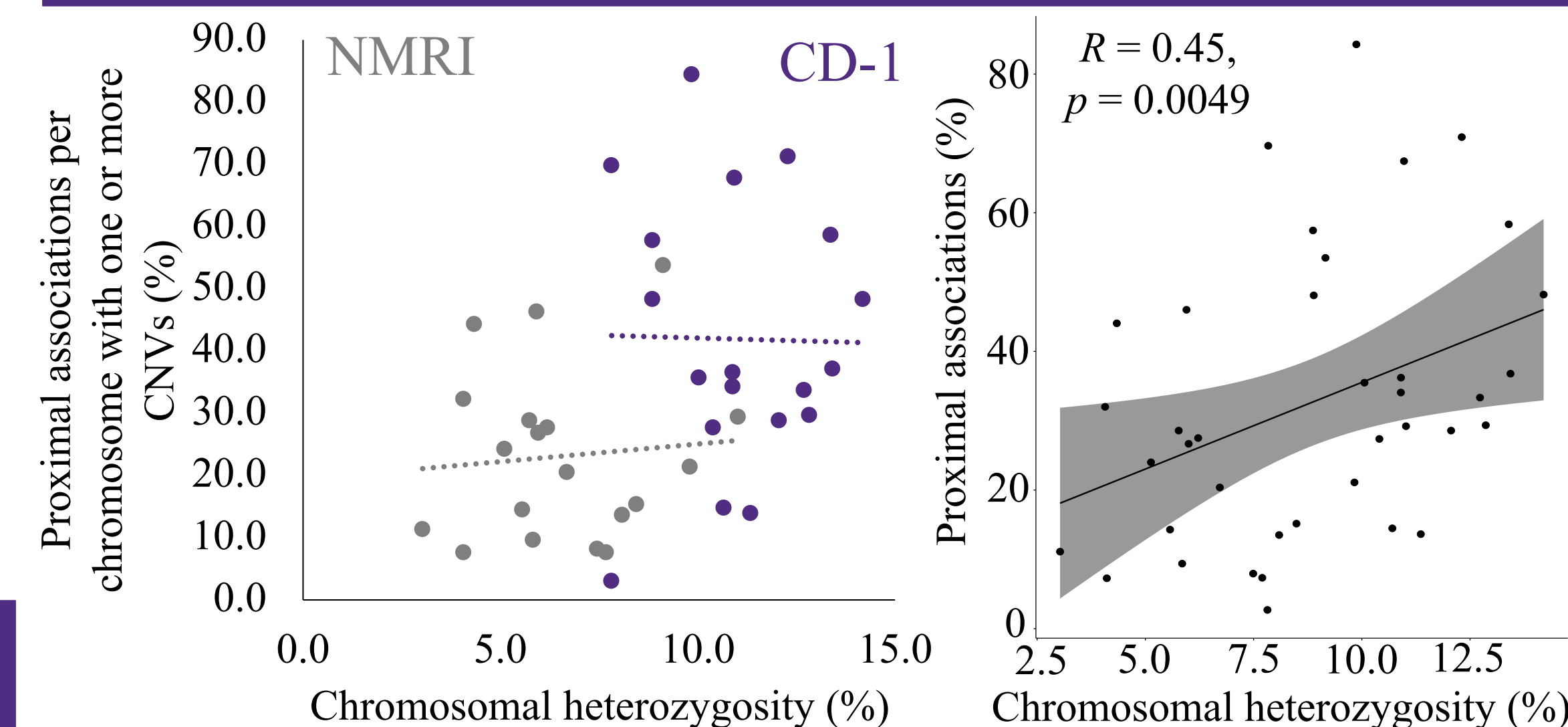
Increased numbers of CNVs with higher levels heterozygosity



CD-1 mice have more chromosomes with proximally distributed heterozygosity



Positive correlation between increasing levels of heterozygosity and proximal association



Conclusions

- Identified a potential association between localized regions of heterozygosity and CNV occurrence
- Future research must determine attributes of heterozygous clusters that are associated with proximal CNVs including density, length and number of heterozygous regions

Acknowledgements