

# R/qtl2 Workshop

---

Karl Broman

Biostatistics and Medical Informatics  
University of Wisconsin – Madison

[kbroman.org/qtl2](http://kbroman.org/qtl2)

[kbroman.org](http://kbroman.org)

[github.com/kbroman](https://github.com/kbroman)

@kwbroman

# Why R/qtl2?

- High-dimensional data  
genotypes and phenotypes
- More diverse crosses  
especially multi-parent populations
- Linear mixed models  
especially in DO/HS/AIL

# R/qtl → R/qtl2

- See [kbroman.org/qtl2/assets/vignettes/rqtl\\_diff.html](http://kbroman.org/qtl2/assets/vignettes/rqtl_diff.html)
- New data file formats
- New data structures
- Split into multiple packages

`qtl2geno, qtl2scan, qtl2plot, qtl2convert`

- New function names

`read.cross()` → `read_cross2()`

`calc.genoprob()` → `calc_genoprob()`

`scanone()` → `scan1()`

- Different treatment of intermediate calculations
- Use of individual IDs for aligning data
- Order of args when subsetting cross objects

`cross[chr,ind]` → `cross2[ind,chr]`

→ R

- `convert2cross2()`
- `summary()`, `n_ind()`, `n_mar()`, ...
- `insert_pseudomarkers()`
- `calc_genoprob()`
- `scan1()`
- `find_peaks()`

# Linear mixed models

$$\begin{aligned} y_i &= \mu + \sum_k \beta_k q_{ik} + \epsilon_i & \epsilon_i &\sim \mathbf{N}(0, \sigma_e^2) \\ &= \mu + \eta_i + \epsilon_i & \eta_i &\sim \mathbf{N}(0, \sigma_p^2) \end{aligned}$$

$$\mathbf{COV}(\eta_i, \eta_j) = \sigma_p^2 (2k_{ij})$$