

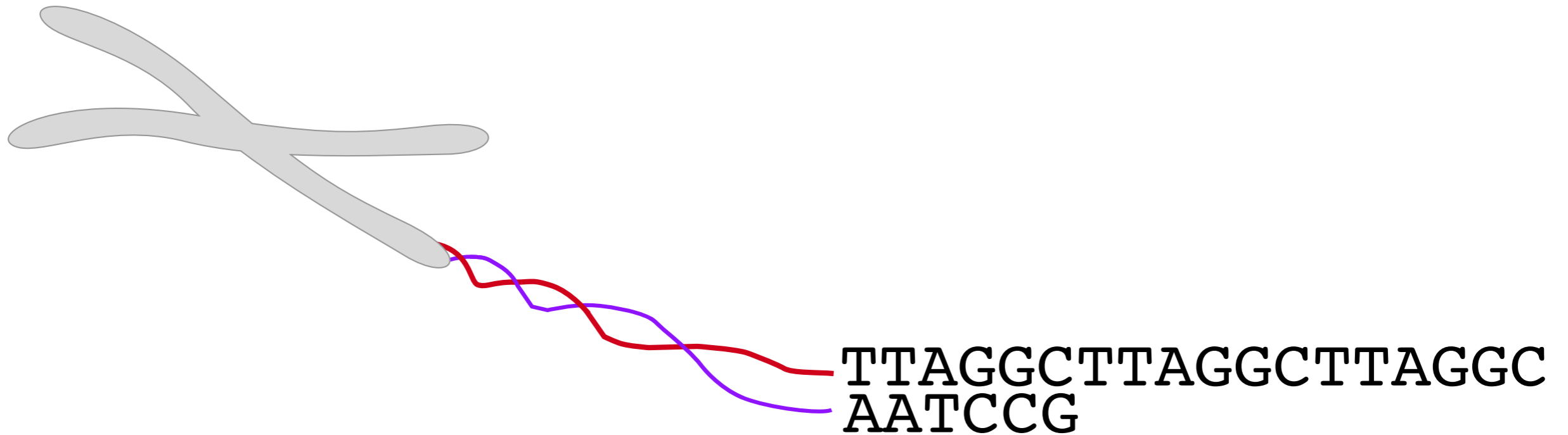
Genome-Wide and Species-Wide Variation in *C. elegans* Reveals
Association of Telomere Length With Population Differences in *pot-2*

Daniel E. Cook

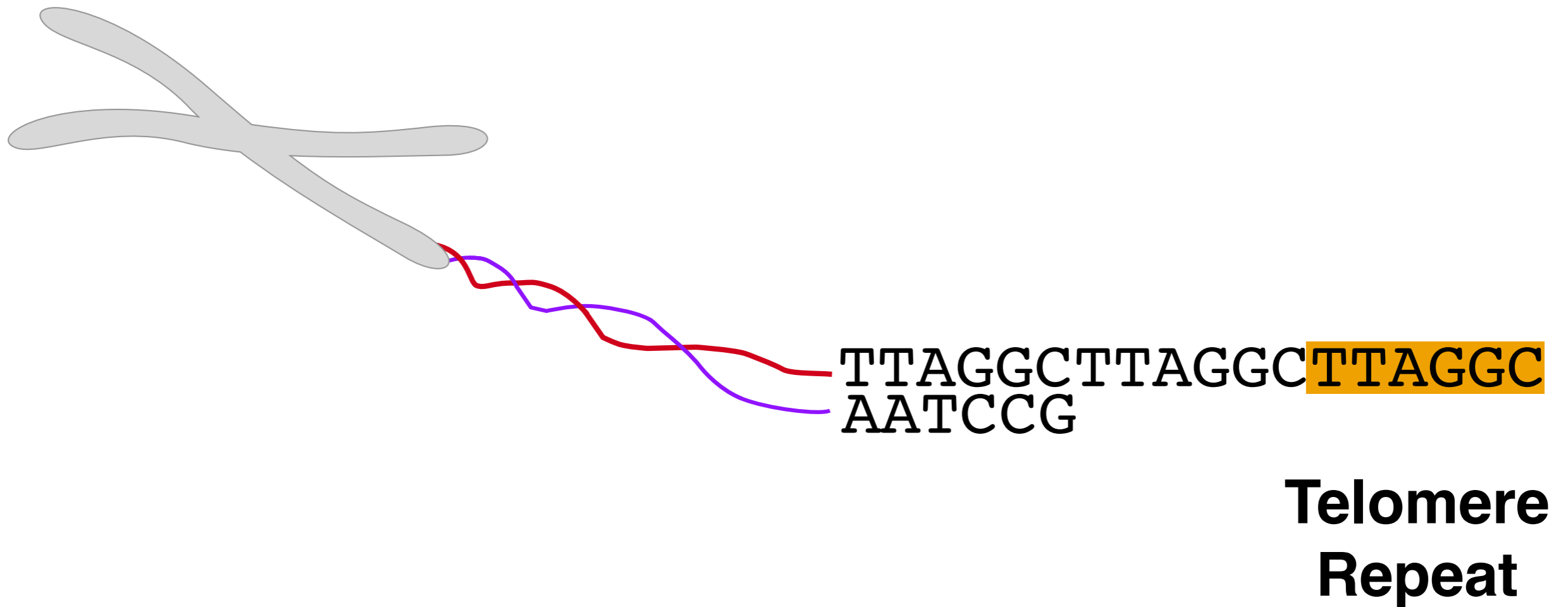
Andersen Lab
Northwestern University
June 25, 2015



Telomeres are located at the terminal ends
of linear chromosomes

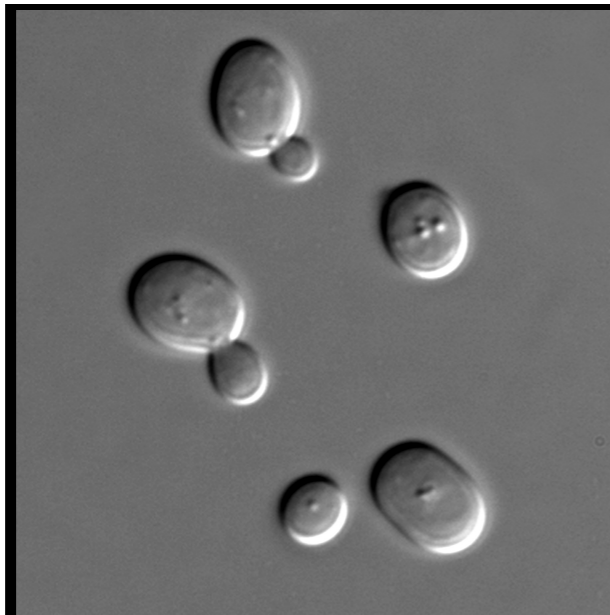


Telomeres are located at the terminal ends
of linear chromosomes



The length of telomeres varies across natural populations

Yeast



Gatbonton, T, *et al.*
PLoS Genetics (2006)

Arabidopsis



Shakirov, E, *et al.*
The Plant Cell (2004)

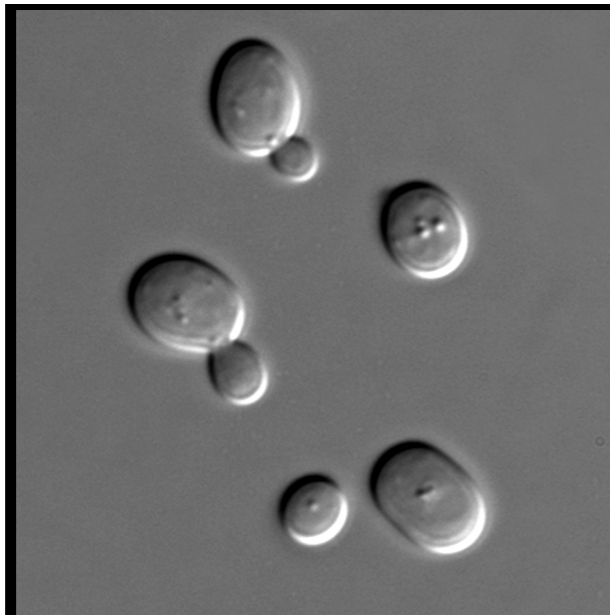
Worms



Raices, M, *et al.*
PLoS Genetics (2005)

Telomeres exhibit natural variation in their length

Yeast



Gatbonton, T, et al.
PLoS Genetics (2006)

Arabidopsis



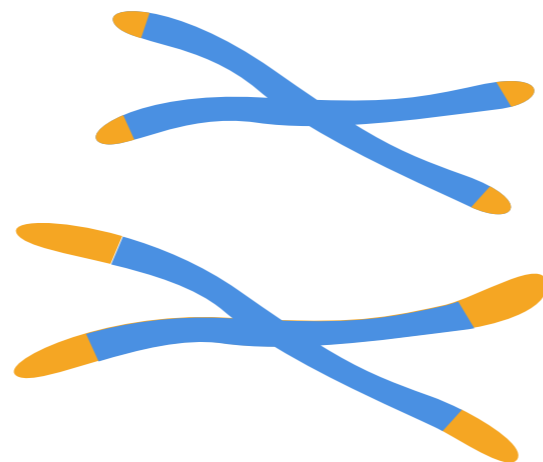
Shakirov, E, et al.
The Plant Cell (2004)

Worms



Raices, M, et al.
PLoS Genetics (2005)

Genetic causes



Fitness consequences

A global collection of 124 *C. elegans* strains



A global collection of 124 *C. elegans* strains



1.6 million single nucleotide polymorphisms (SNPs)

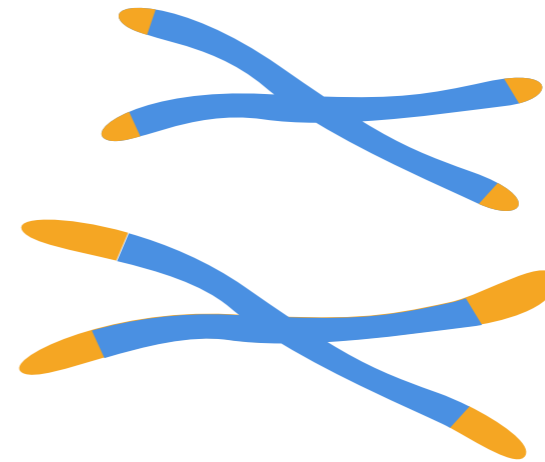
Correlation between genotype and phenotype identifies associated regions of the genome



TAAAGCTTATCAGTAT
TGTTTCCTGGAGAGTAG
GAGGGGTTTAAACGATG
TGGGATTTTCTGTIG
TTAAATCCTGGAACAC
AACTAGCCGCATTACA

1.6 Million
SNPs

**Genetic
variation**



**Telomere length
variation**

Illumina short-read sequence data can be used to estimate telomere length

Non-telomeric reads possess few or no telomeric repeats

GC GGCTGA GTTTAG GTTTAG ATGCGT **TTAGGC** GTGTGT TTGTGT AAAGTG TT

Illumina short-read sequence data can be used to estimate telomere length

Non-telomeric reads possess few or no telomeric repeats

GC GGCTGA GTTTAG GTTTAG ATGCGT **TTAGGC** GTGTGT TTGTGT AAAGTG TT

Telomeric reads contain many telomeric repeats

GC **TTAGGC** **TTAGGC** **TTAGGC** **TTAGGC** **TTAGGC** **TTAGGC** **TTAGGC** **TTAGGC** **TTAGGC** TT

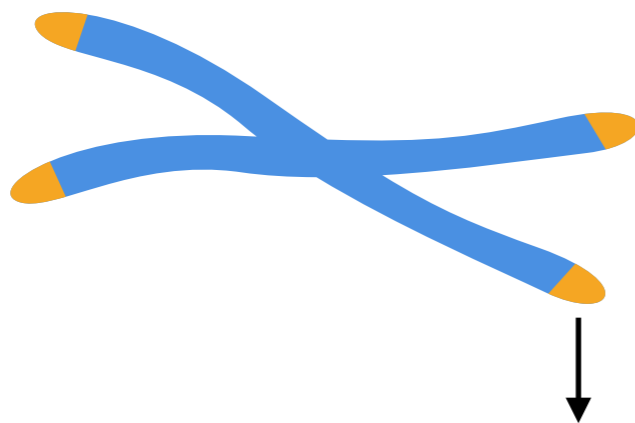
Illumina short-read sequence data can be used to estimate telomere length

Non-telomeric reads possess few or no telomeric repeats

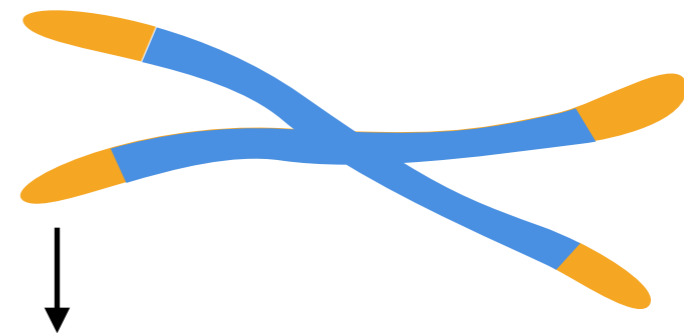
GC GGCTGA GTTTAG GTTTAG ATGCGT **TTAGGC** GTGTGT TTGTGT AAAGTG TT

Telomeric reads contain many telomeric repeats

GC **TTAGGC** **TTAGGC** **TTAGGC** **TTAGGC** **TTAGGC** **TTAGGC** **TTAGGC** **TTAGGC** **TTAGGC** TT

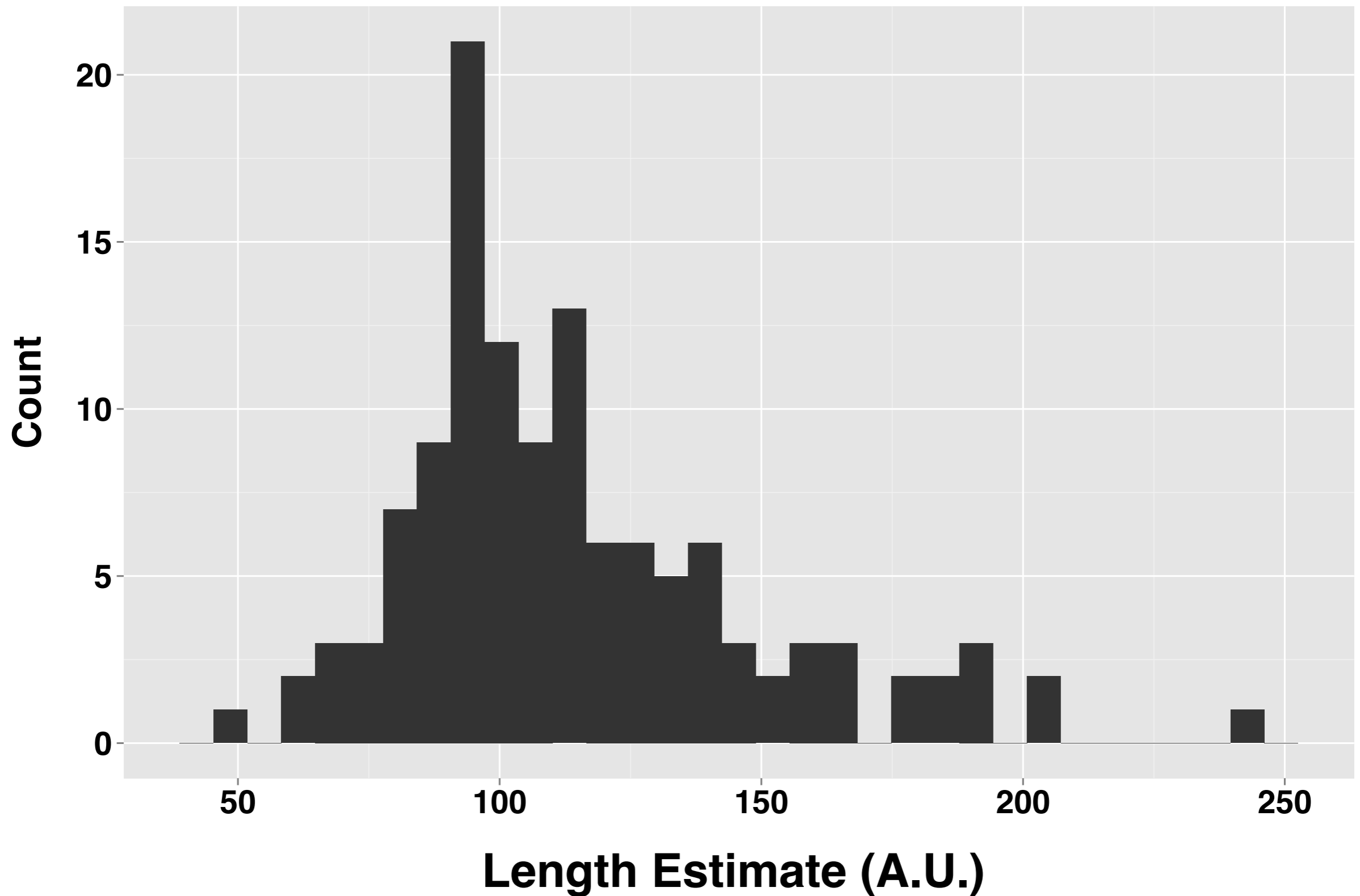


Low proportion
of telomeric reads



High proportion
of telomeric reads

Natural variation of telomere length in *C. elegans* wild isolates



Correlation between genotype and phenotype identifies associated regions of the genome



TAAAGCTTATCAGTAT
TGTTTCCTGGAGAGTAG
GAGGGGGAAGGATG
TGGGATTTTCTGTIG
TTAAATCCTGGAACAC
AACTAGCCGCATTACA

1.6 Million
SNPs

**Genetic
variation**

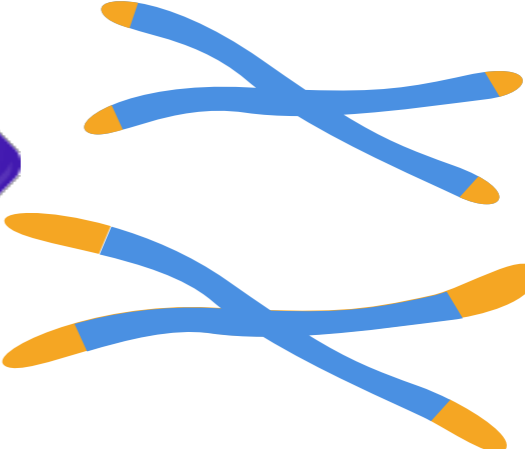



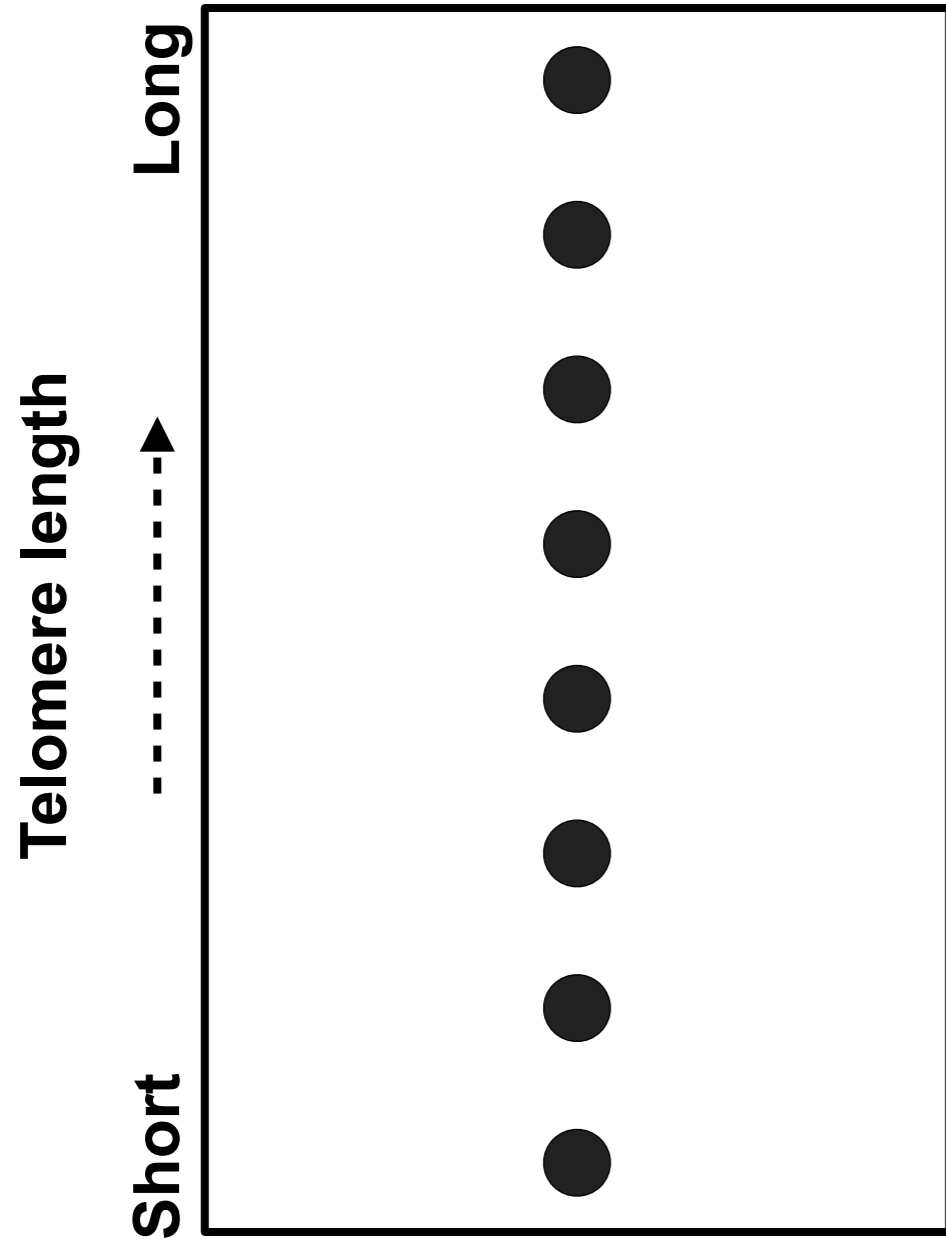
Diagram illustrating telomere length variation, showing blue chromosome arms with yellow tips (telomeres).

**Telomere length
variation**

Associated genes

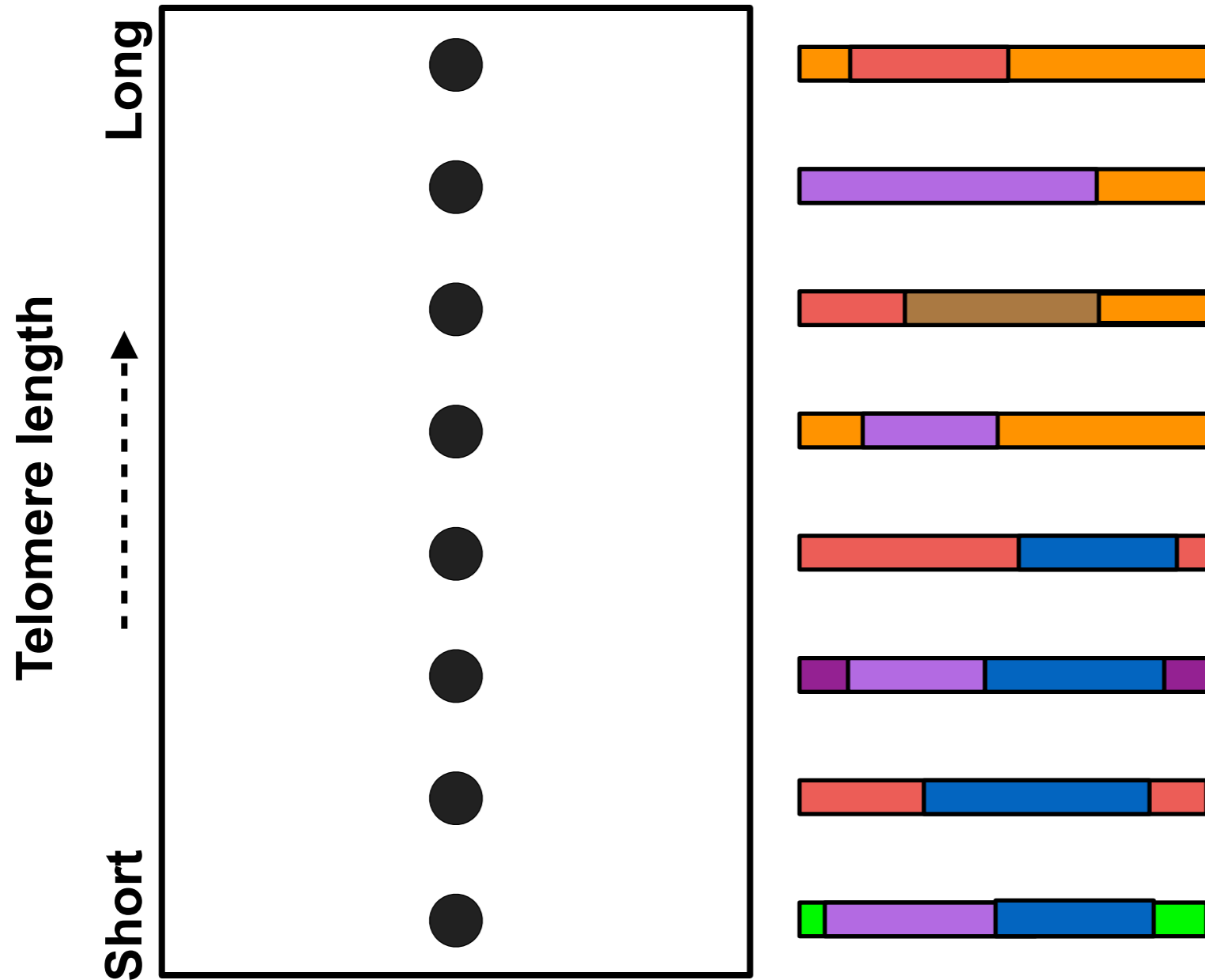
Association mapping:

Correlating genotype with phenotype using wild strains



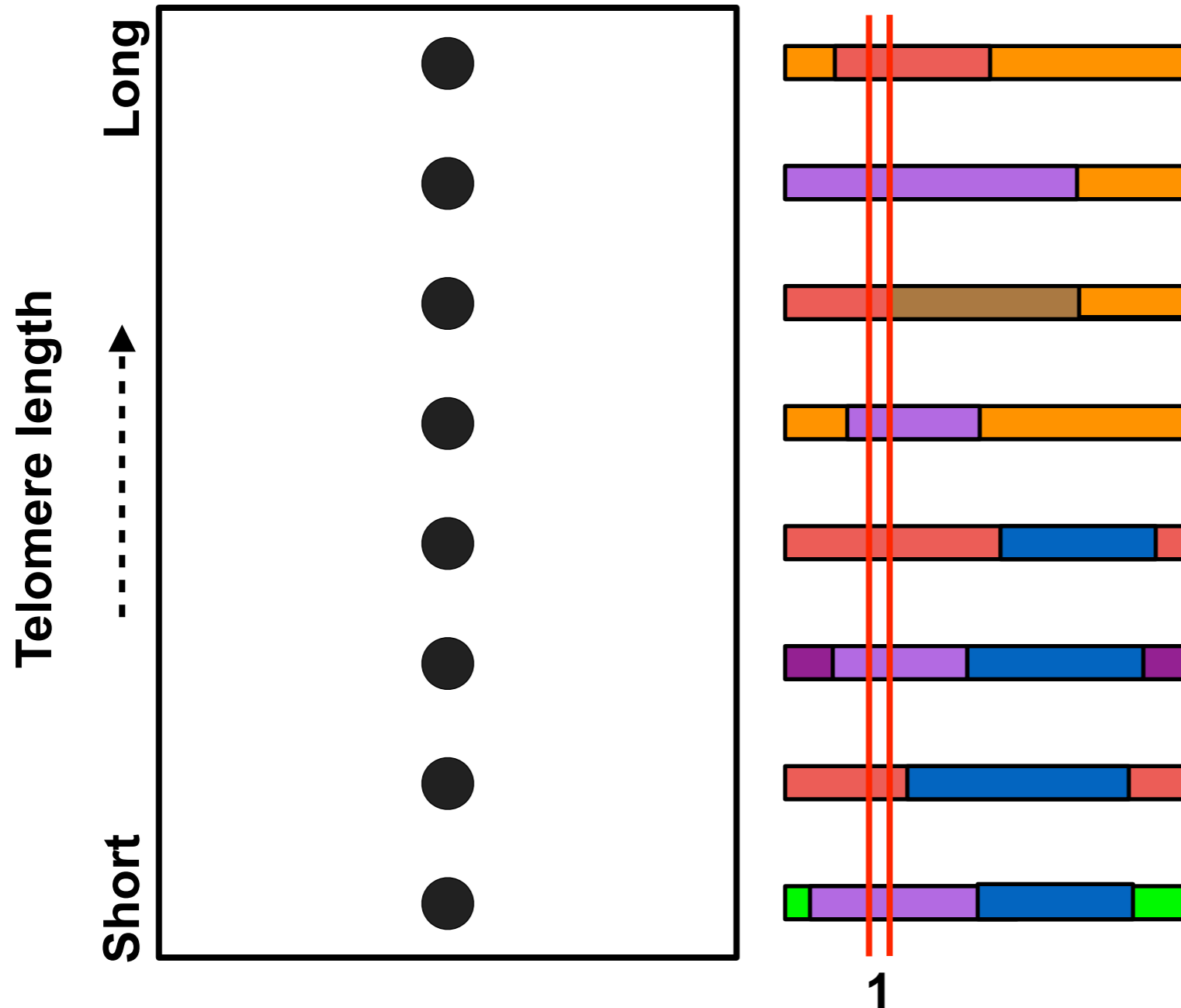
Association mapping:

Correlating genotype with phenotype using wild strains



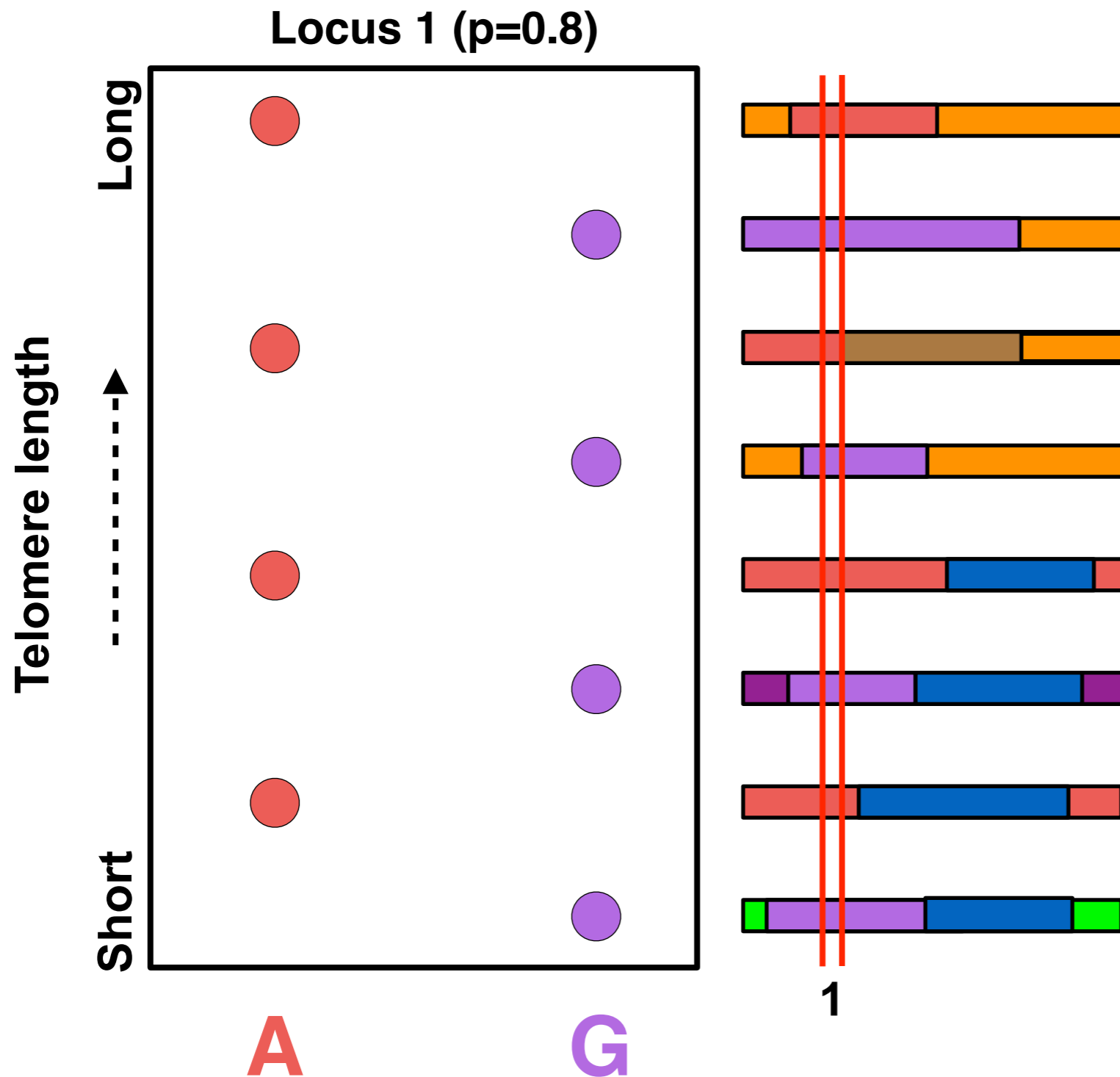
Association mapping:

Correlating genotype with phenotype using wild strains



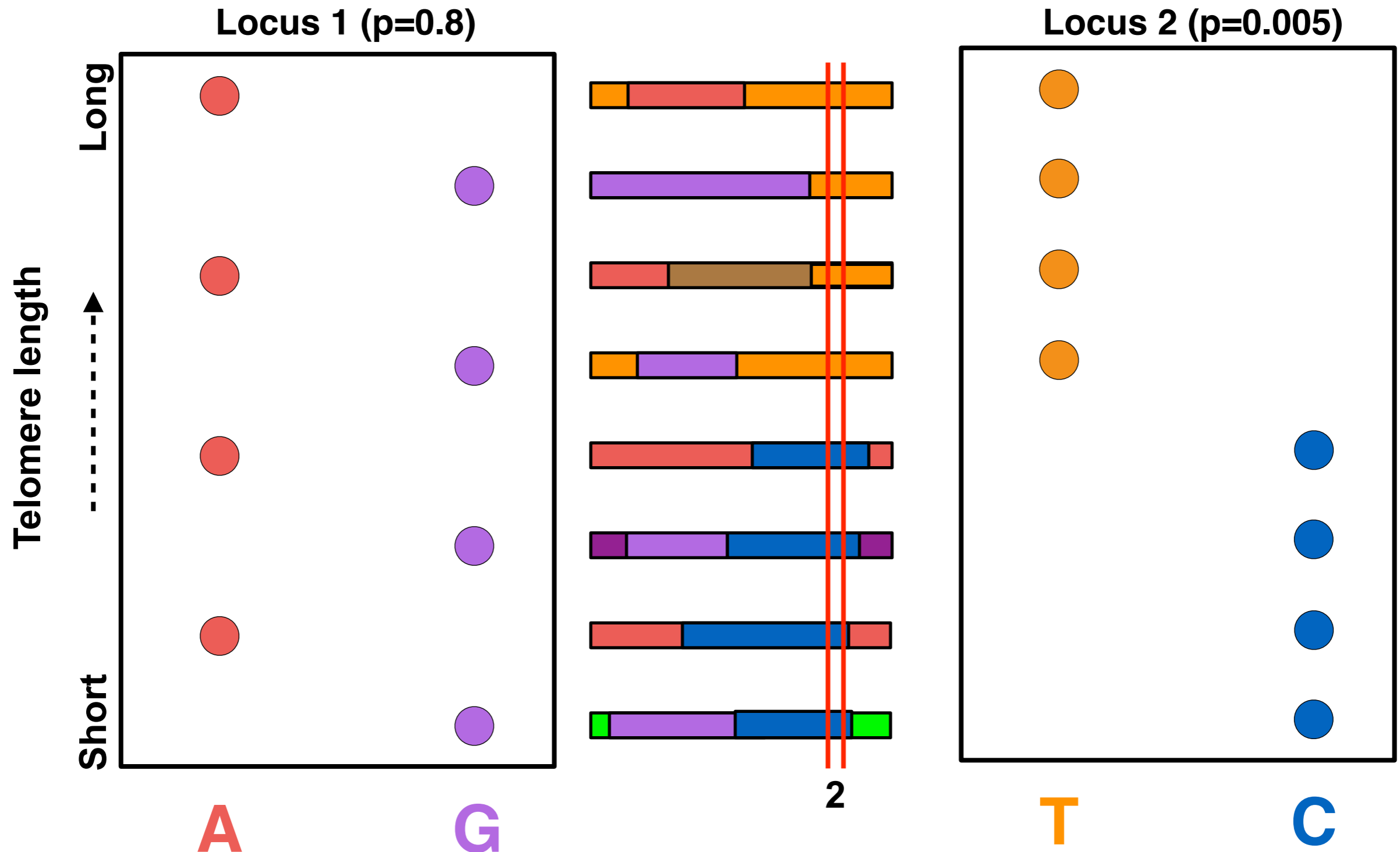
Association mapping:

Correlating genotype with phenotype using wild strains



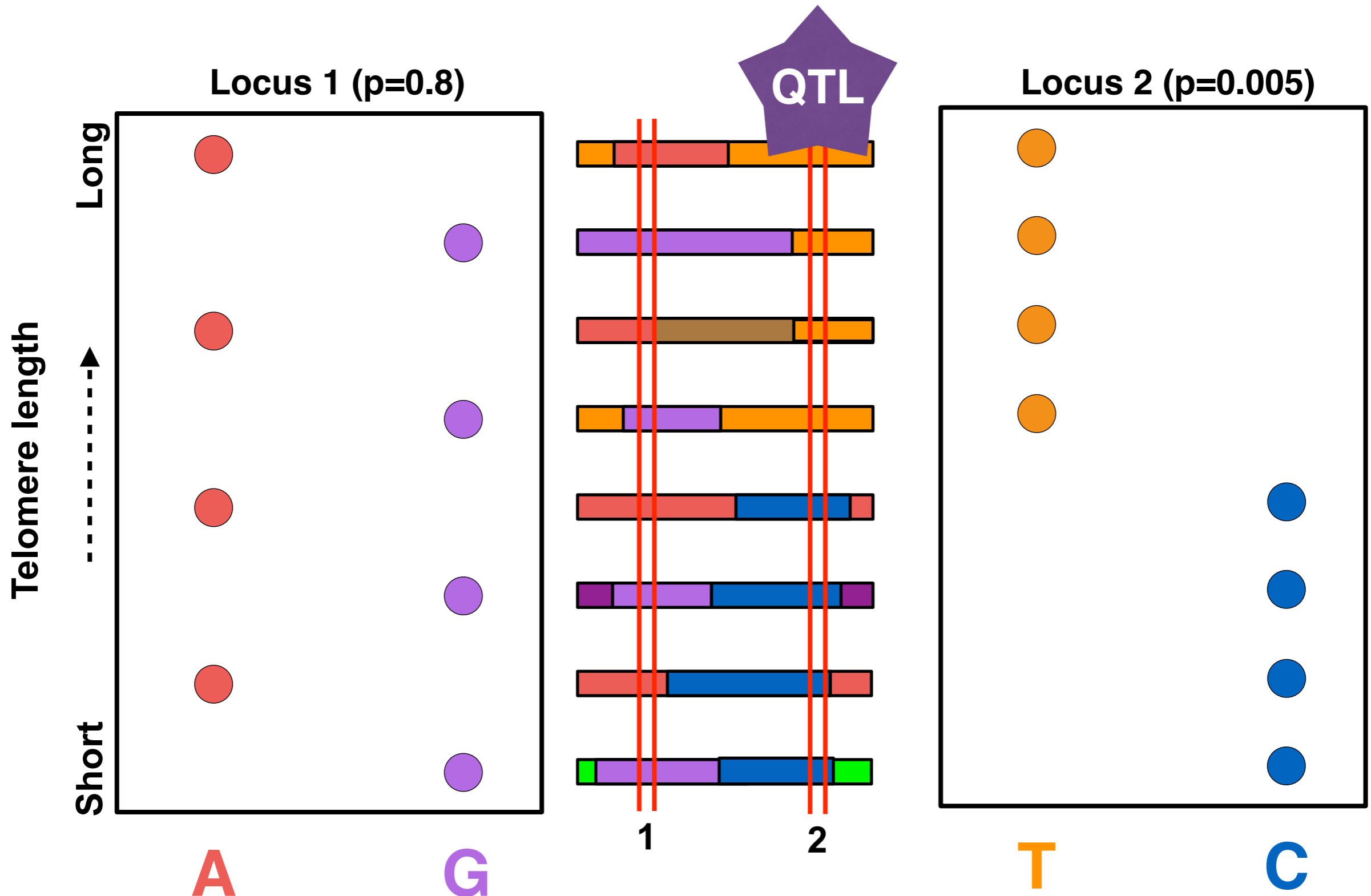
Association mapping:

Correlating genotype with phenotype using wild strains

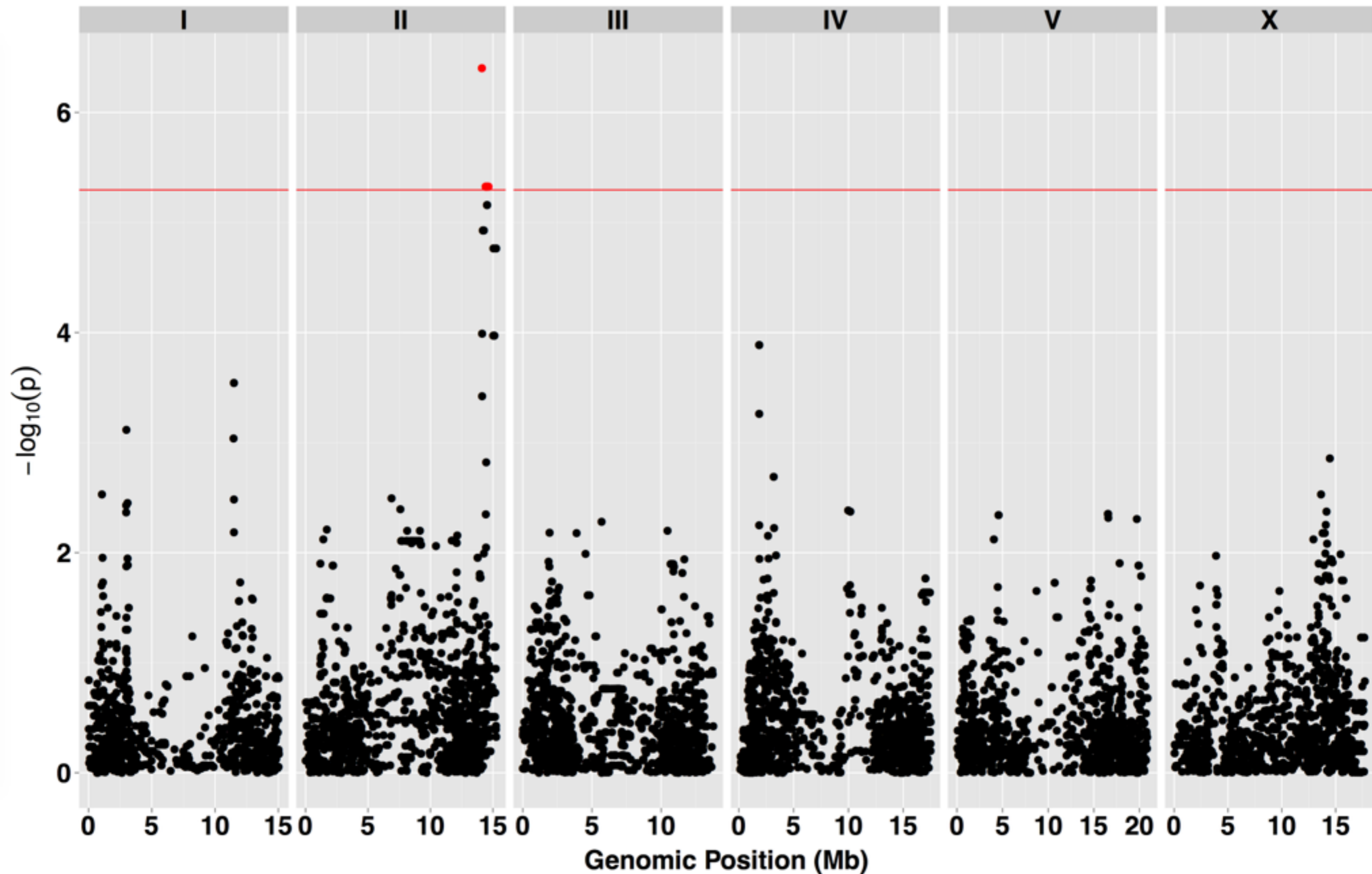


Association mapping:

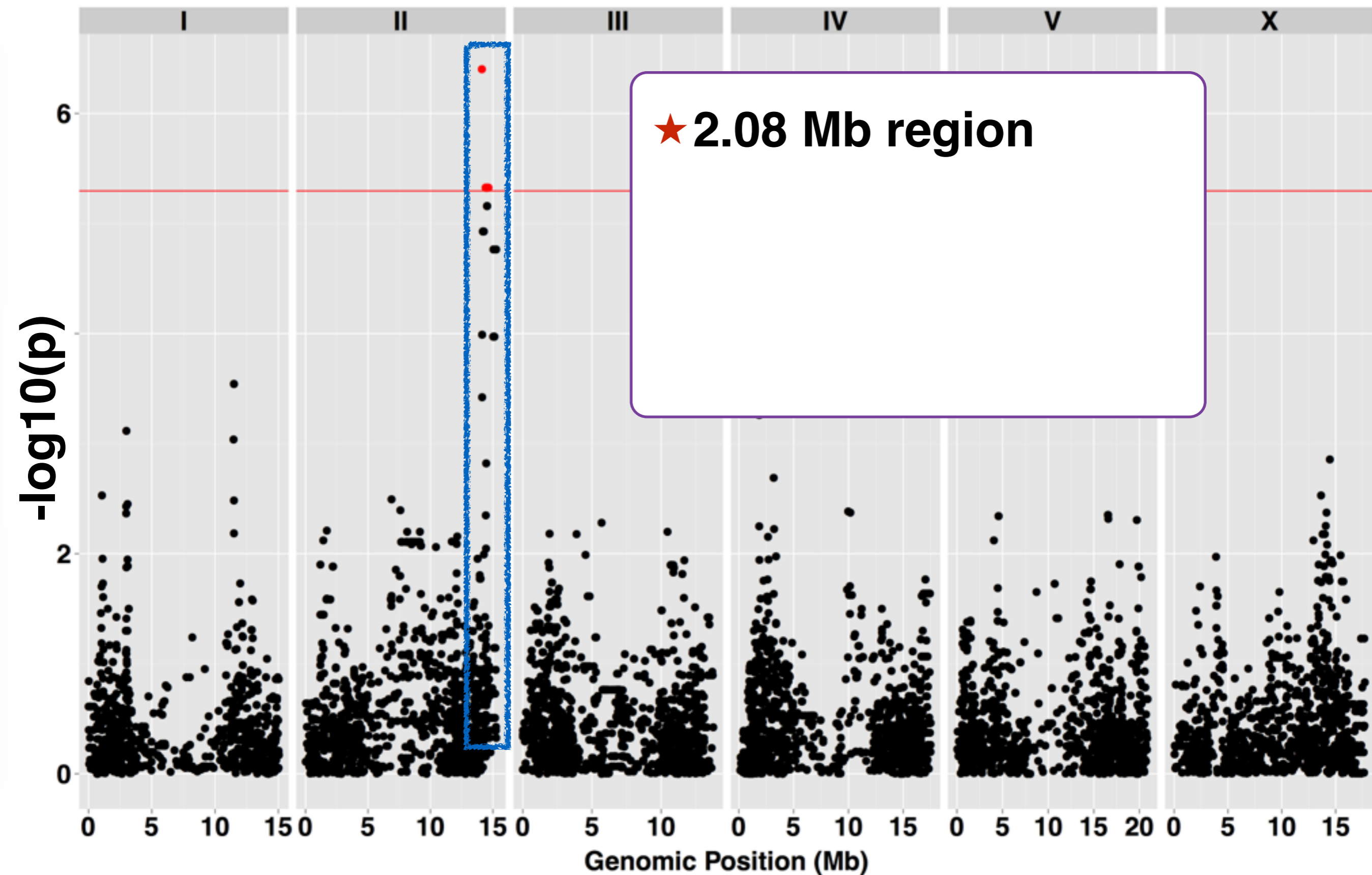
Correlating genotype with phenotype using wild strains



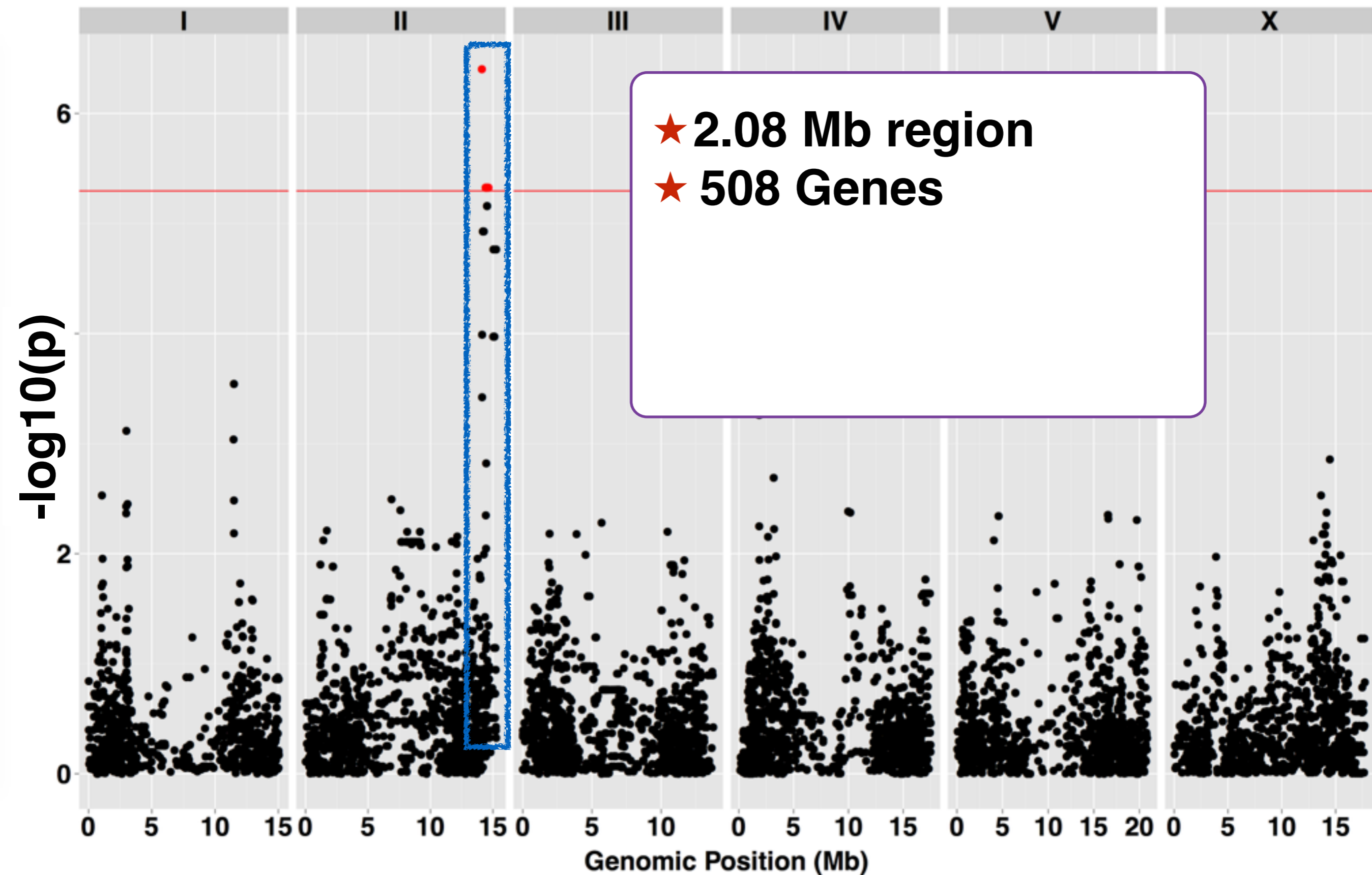
Genome-wide association identifies a region on the right arm of chromosome II associated with telomere length



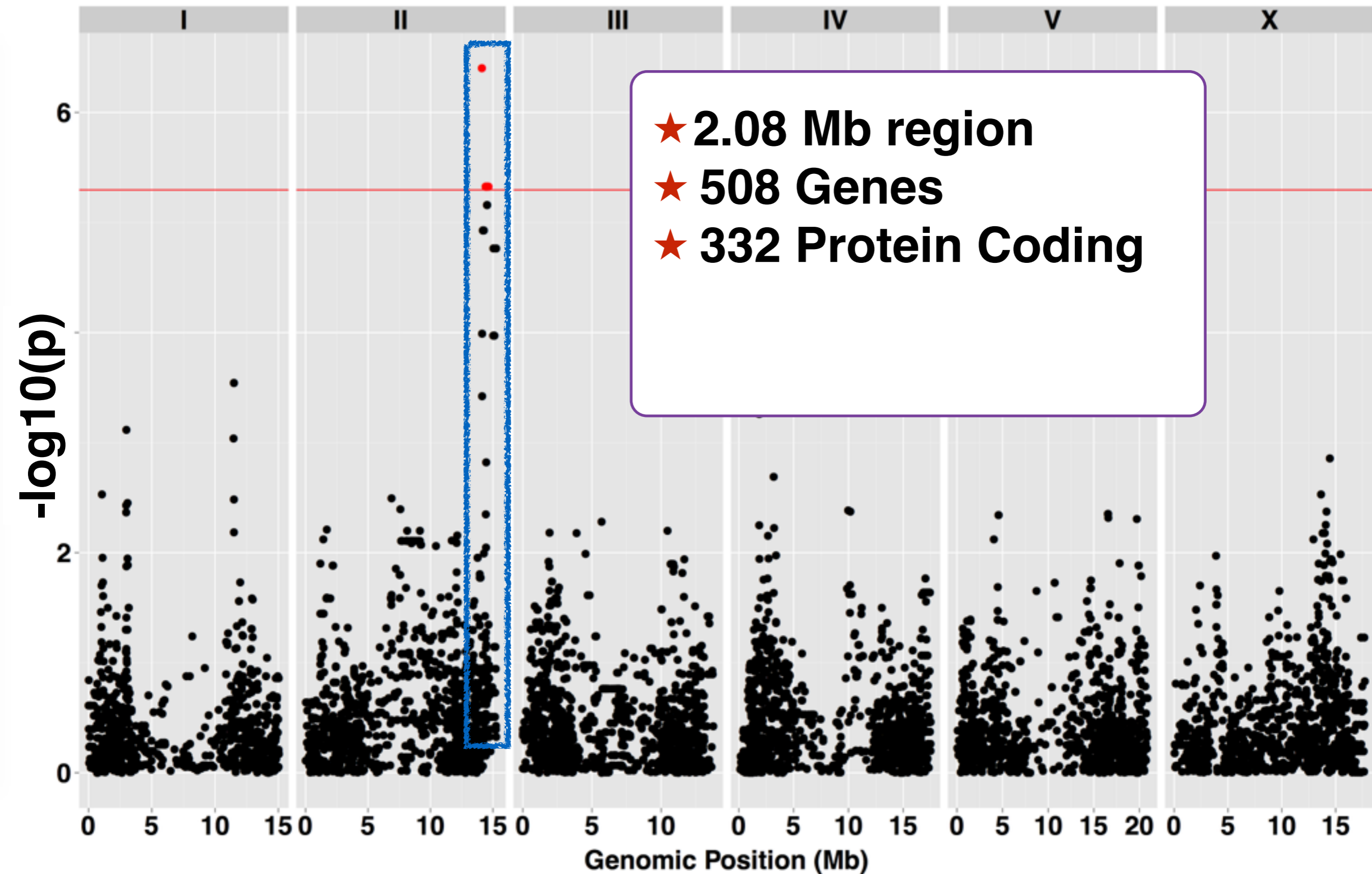
Genome-wide association identifies a region on the right arm of chromosome II associated with telomere length



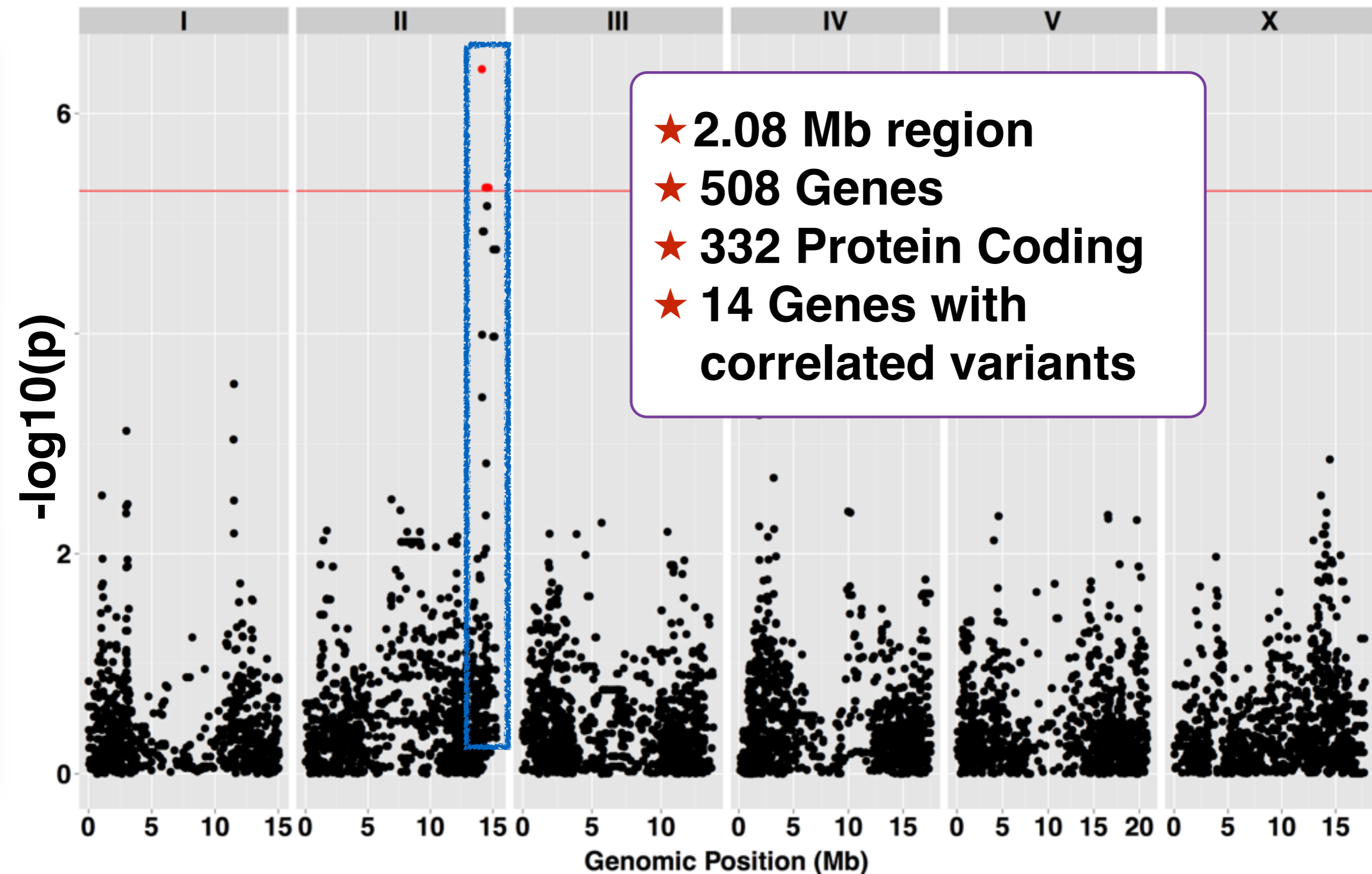
Genome-wide association identifies a region on the right arm of chromosome II associated with telomere length



Genome-wide association identifies a region on the right arm of chromosome II associated with telomere length



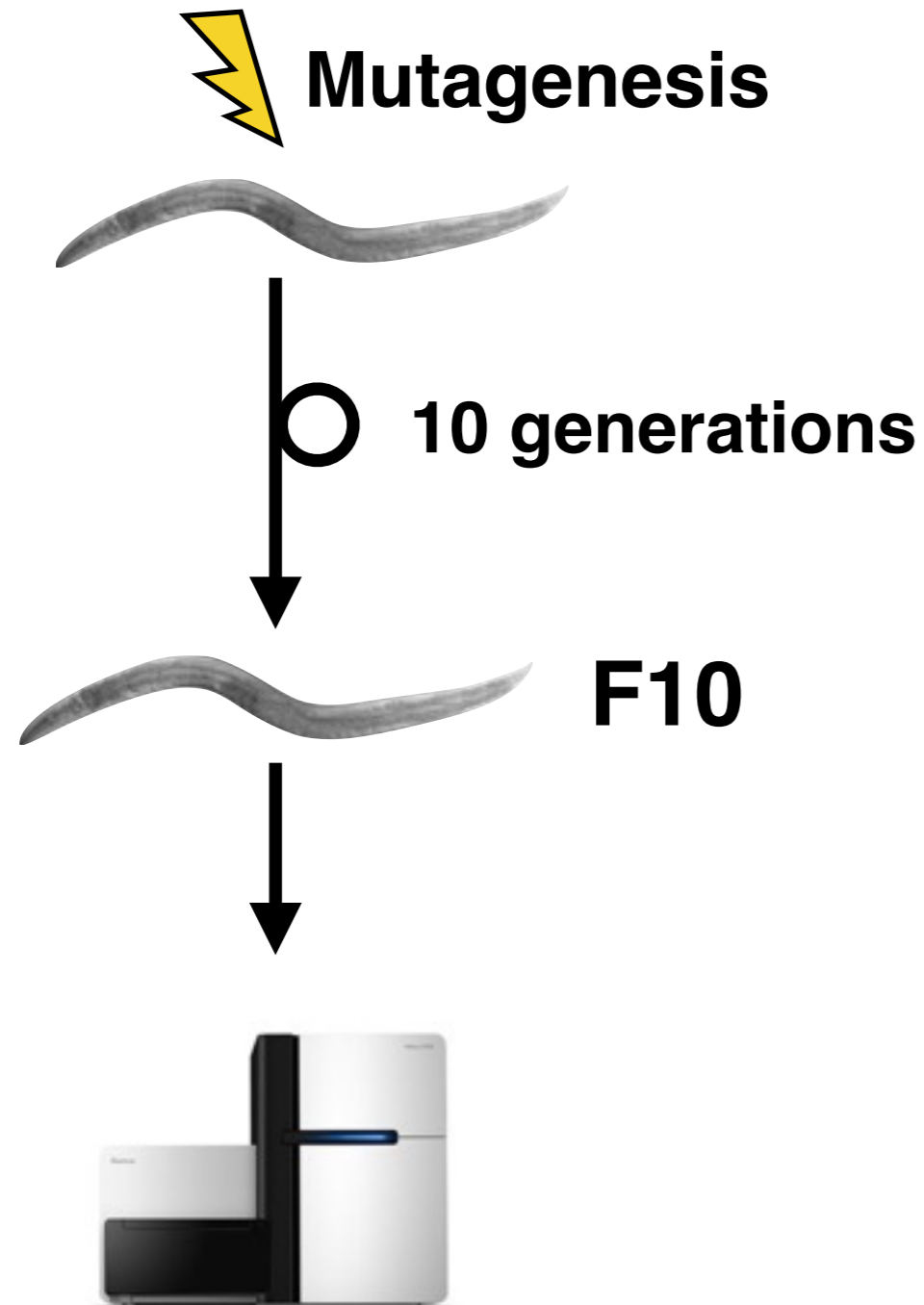
Genome-wide association identifies a region on the right arm of chromosome II associated with telomere length



14 genes are maximally correlated with telomere length

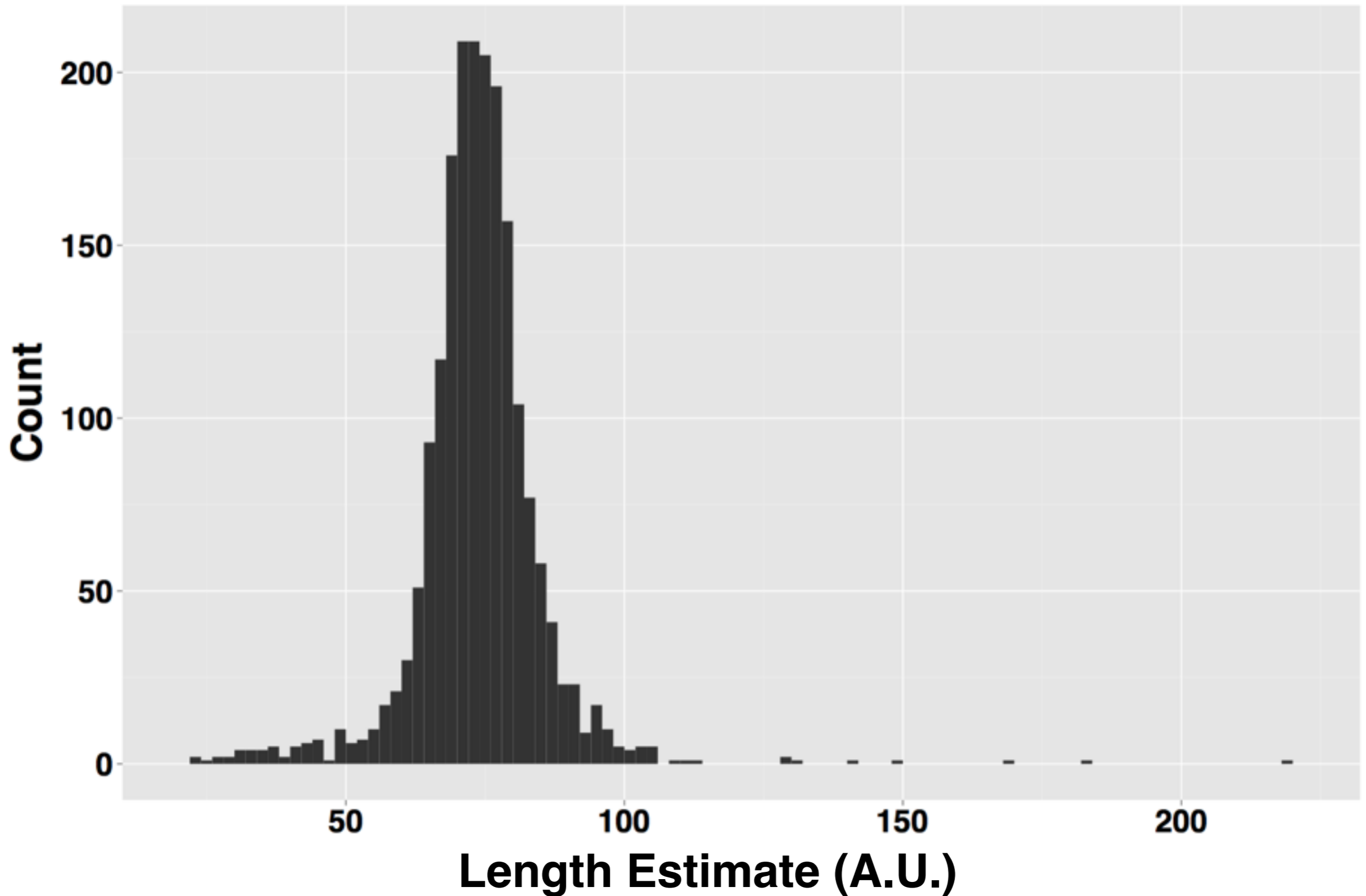
- *C04H5.7*
- *C13B4.1*
- *K09E4.2*
- *Y54G11A.14*
- *ace-4*
- *btb-19*
- *btb-20*
- *dmsr-6*
- *gcy-29*
- *mltn-9*
- *pot-2*
- *srh-42*
- *unc-52*
- *wee-1*

Million Mutation Project (MMP) data can be used to identify genes that regulate telomere length

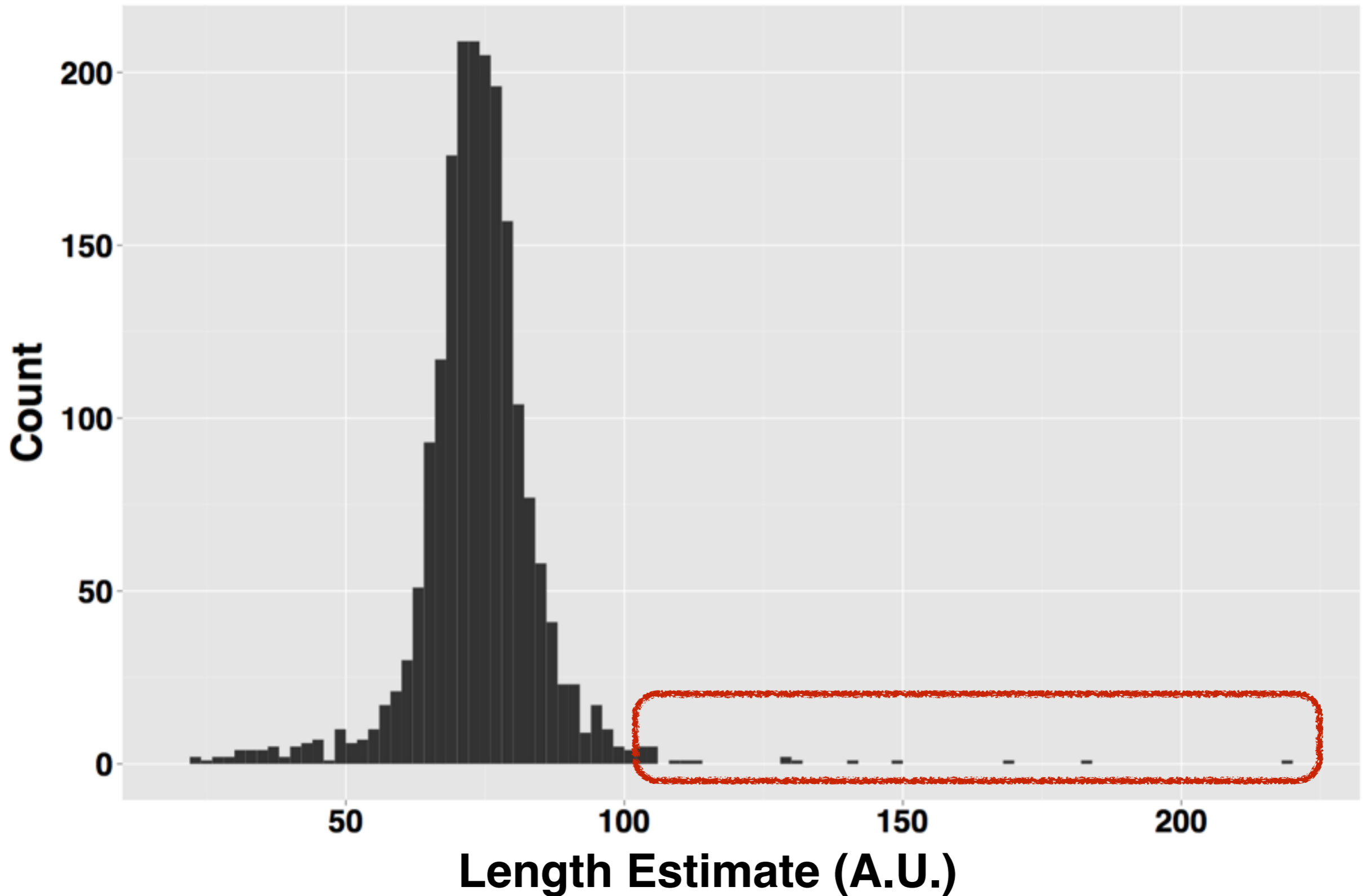


- Mutagenized N2 Background
- ~2000 mutant strains
- Maintained for ten generations
- Whole-genome sequenced

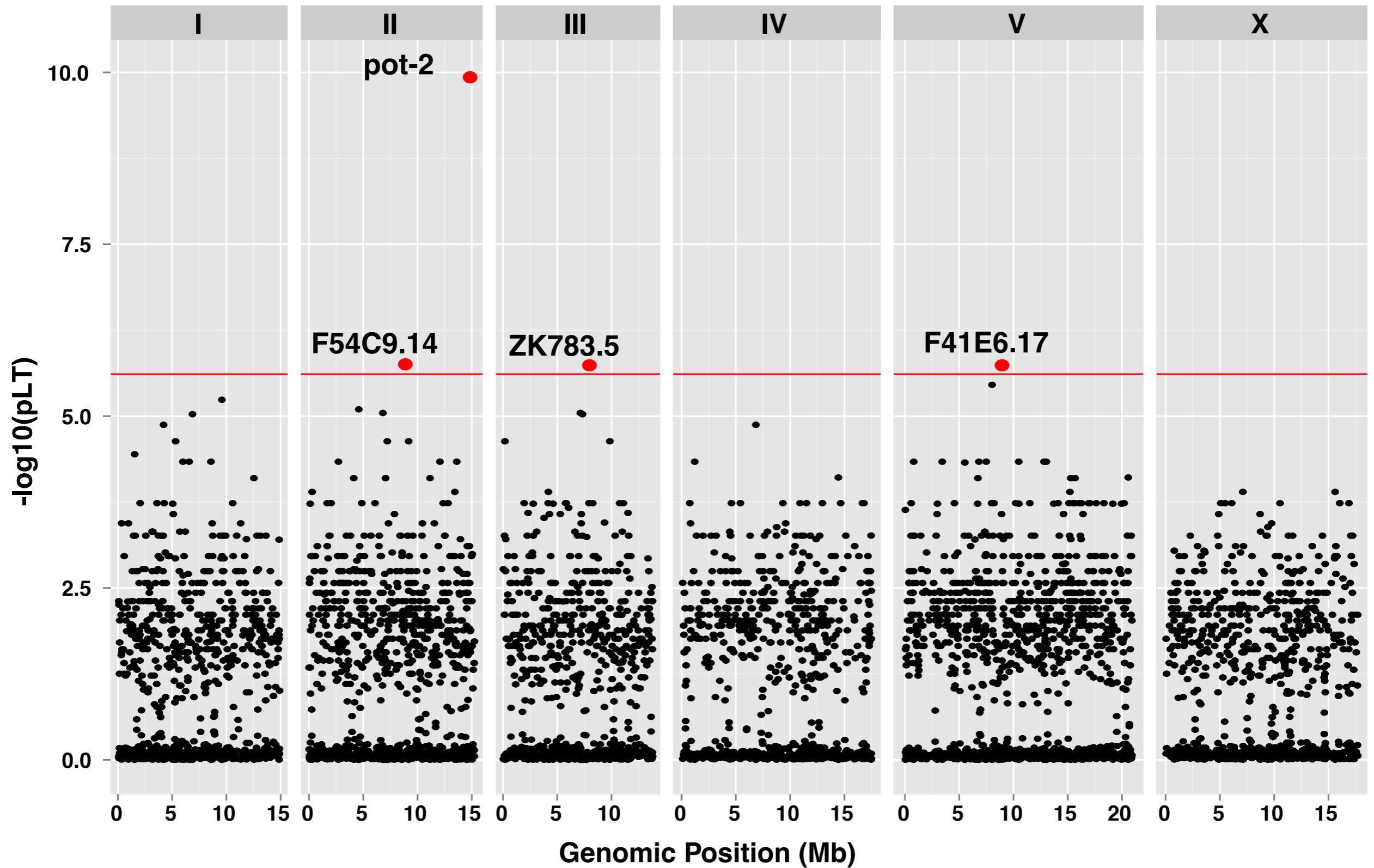
A subset of MMP strains have excessively long telomeres



**A subset of MMP strains
have excessively long telomeres**



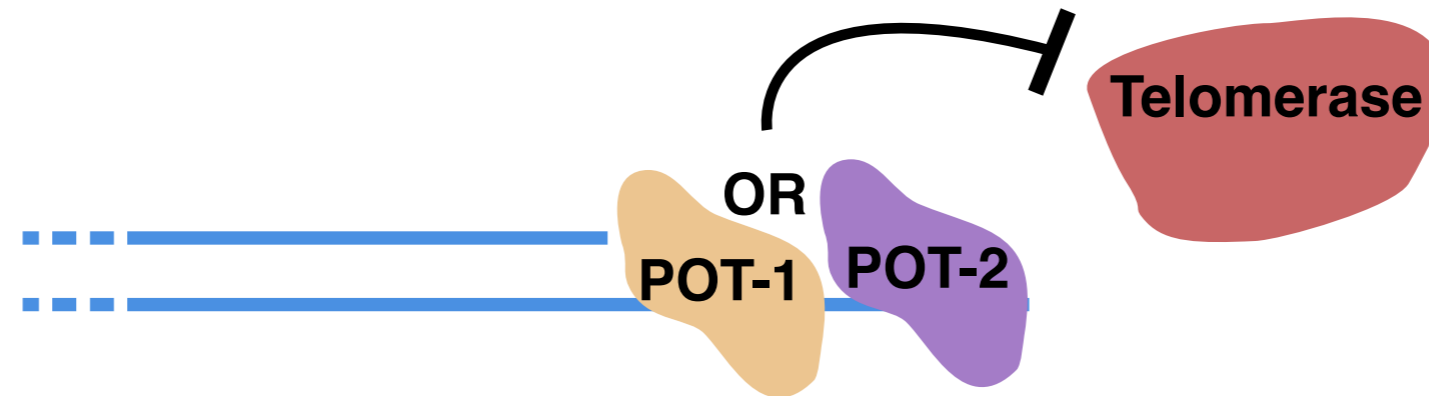
Long telomere strains are enriched for *pot-2* mutations



Two orthogonal techniques indicate that *pot-2* contributes to regulation of telomere length

- *C04H5.7*
- *C13B4.1*
- *K09E4.2*
- *Y54G11A.14*
- *ace-4*
- *btb-19*
- *btb-20*
- *dmsr-6*
- *gcy-29*
- *mltn-9*
- *pot-2*
- *srh-42*
- *unc-52*
- *wee-1*

pot-2 (Protection of Telomeres 2) inhibits telomerase activity



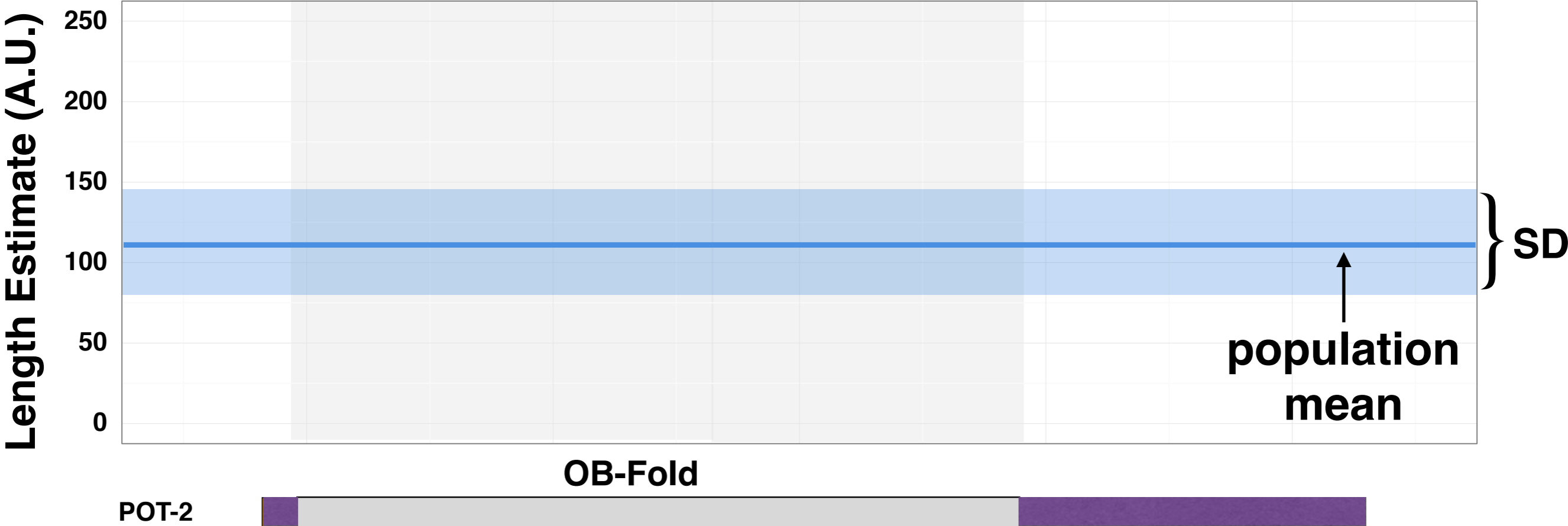
- Binds single-stranded ends of telomeres
- Loss of *pot-2* causes lengthening of telomeres

De Lange, T
Genes and Development (2005)

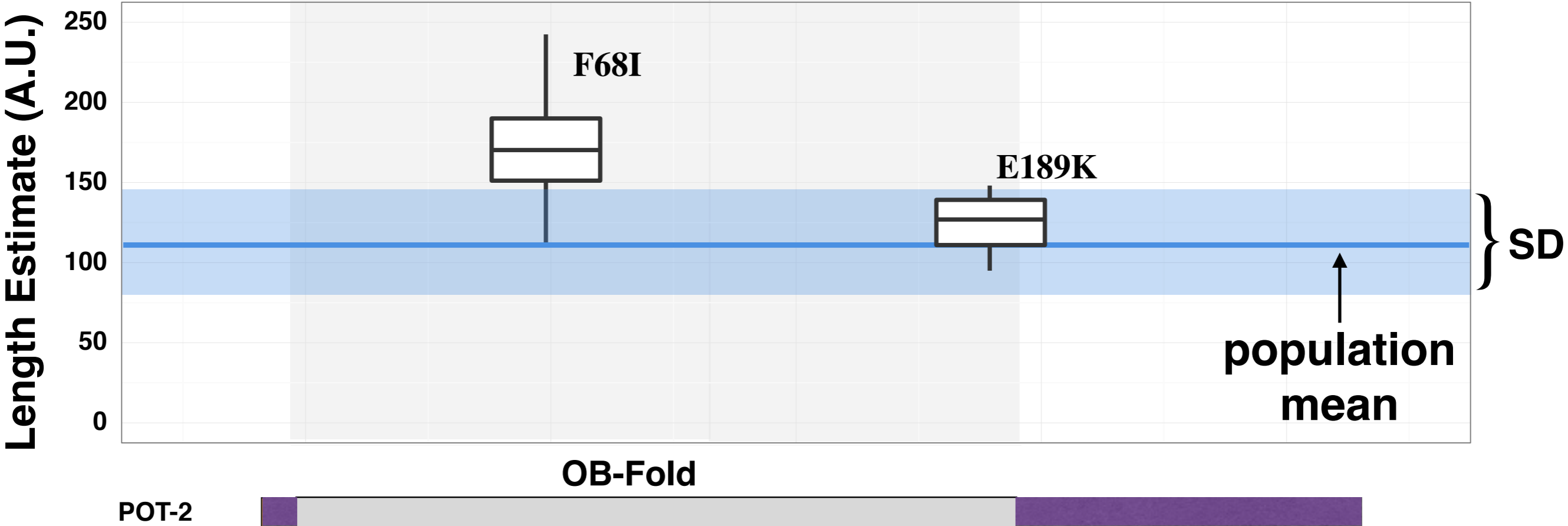
Raices, M. *et al.*
Cell (2008).

Shtessel, L., *et al.*
G3 (2013).

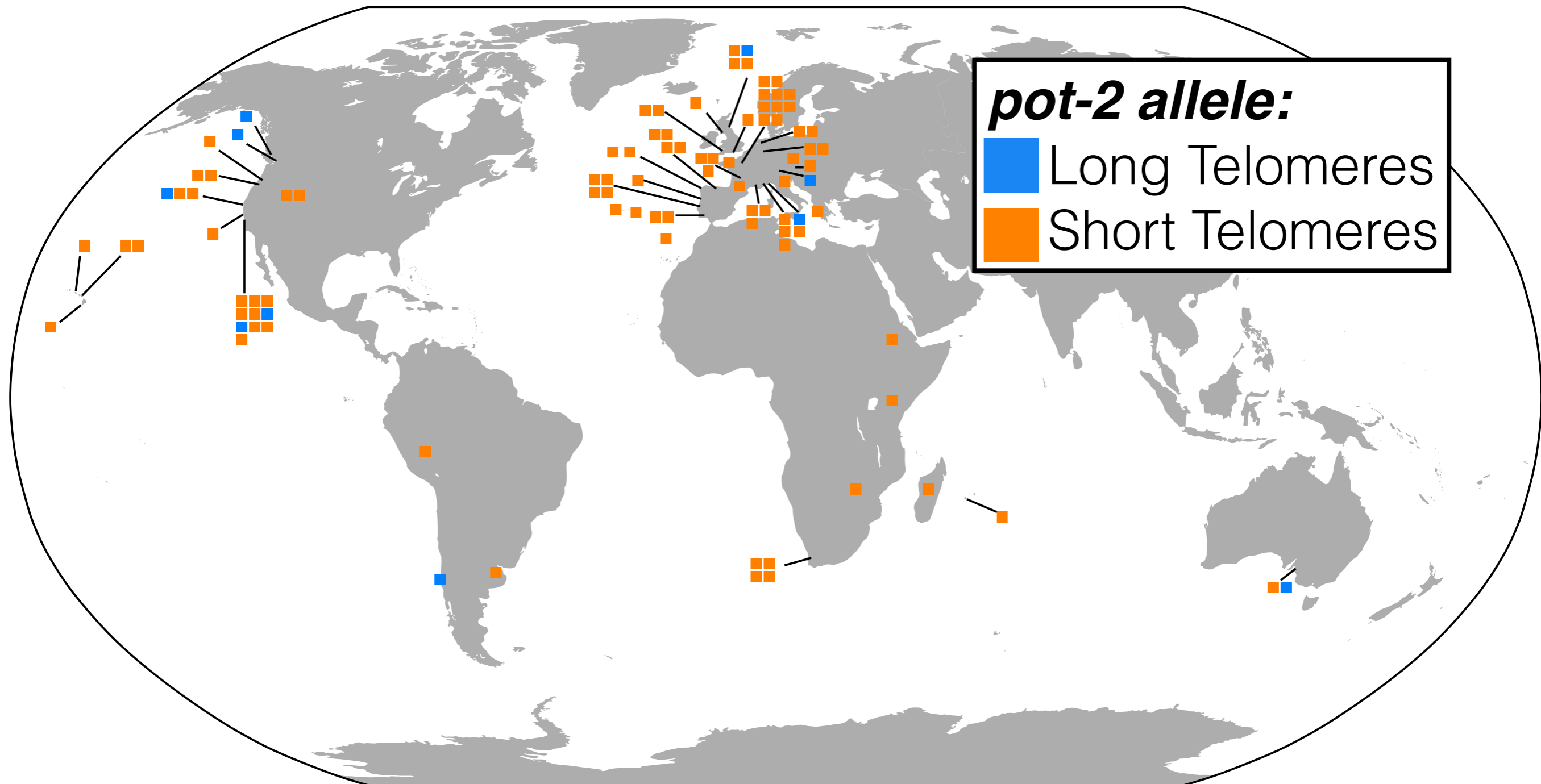
Variation in the OB-fold of POT-2 is associated with longer telomere lengths



Variation in the OB-fold of POT-2 is associated with longer telomere lengths



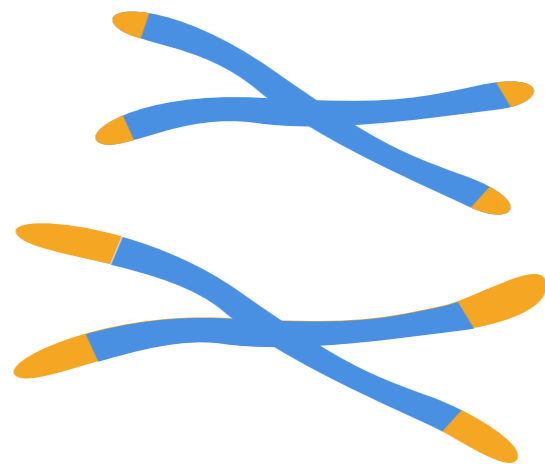
The distribution of the long-telomere allele is not geographically restricted



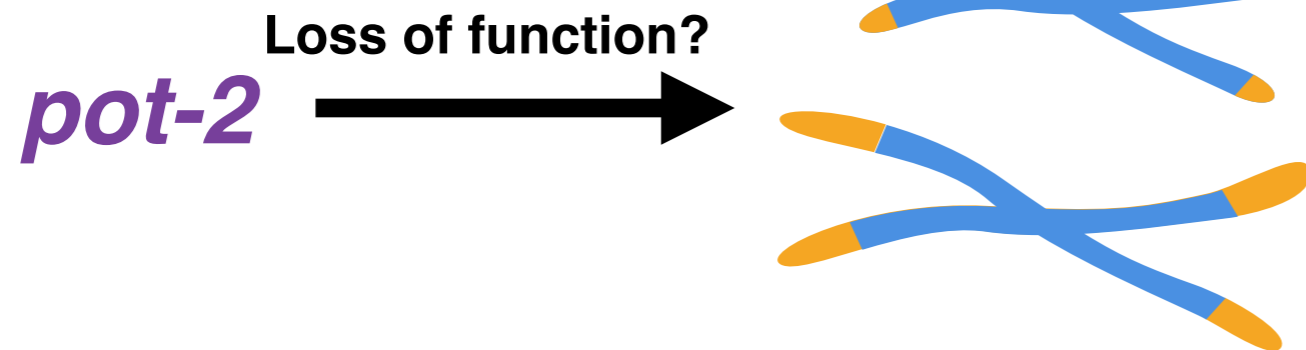
Telomeres exhibit natural variation in their length



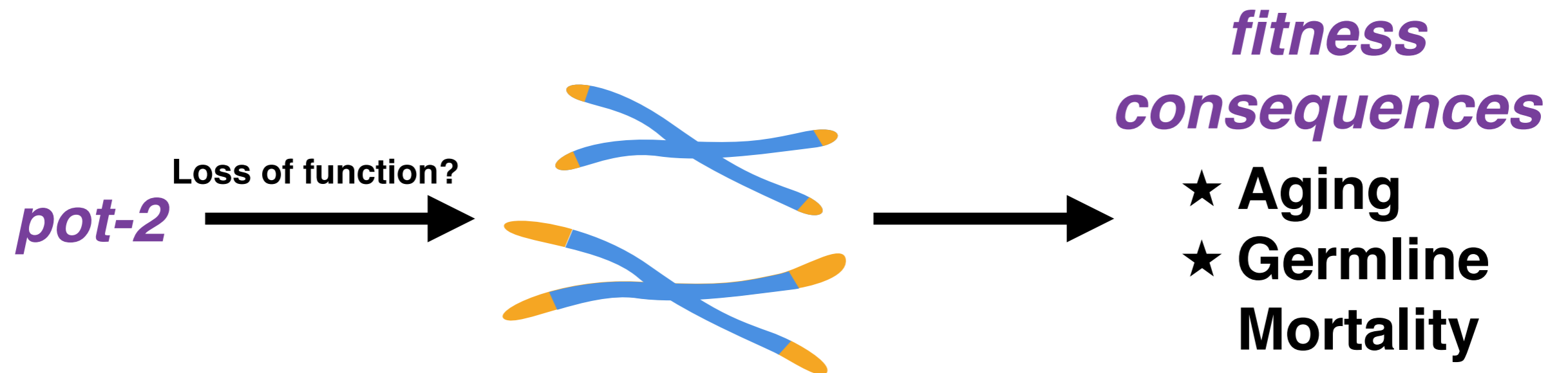
***Genetic
Causes***

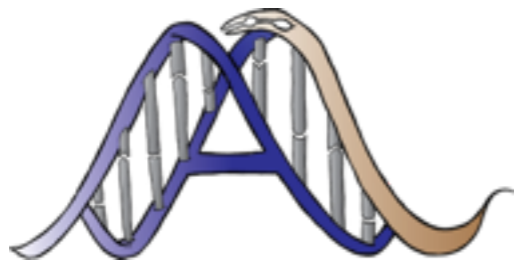


Variation in *pot-2* is likely contributing to differences observed in telomere length.



Natural variation in telomere length may have fitness consequences

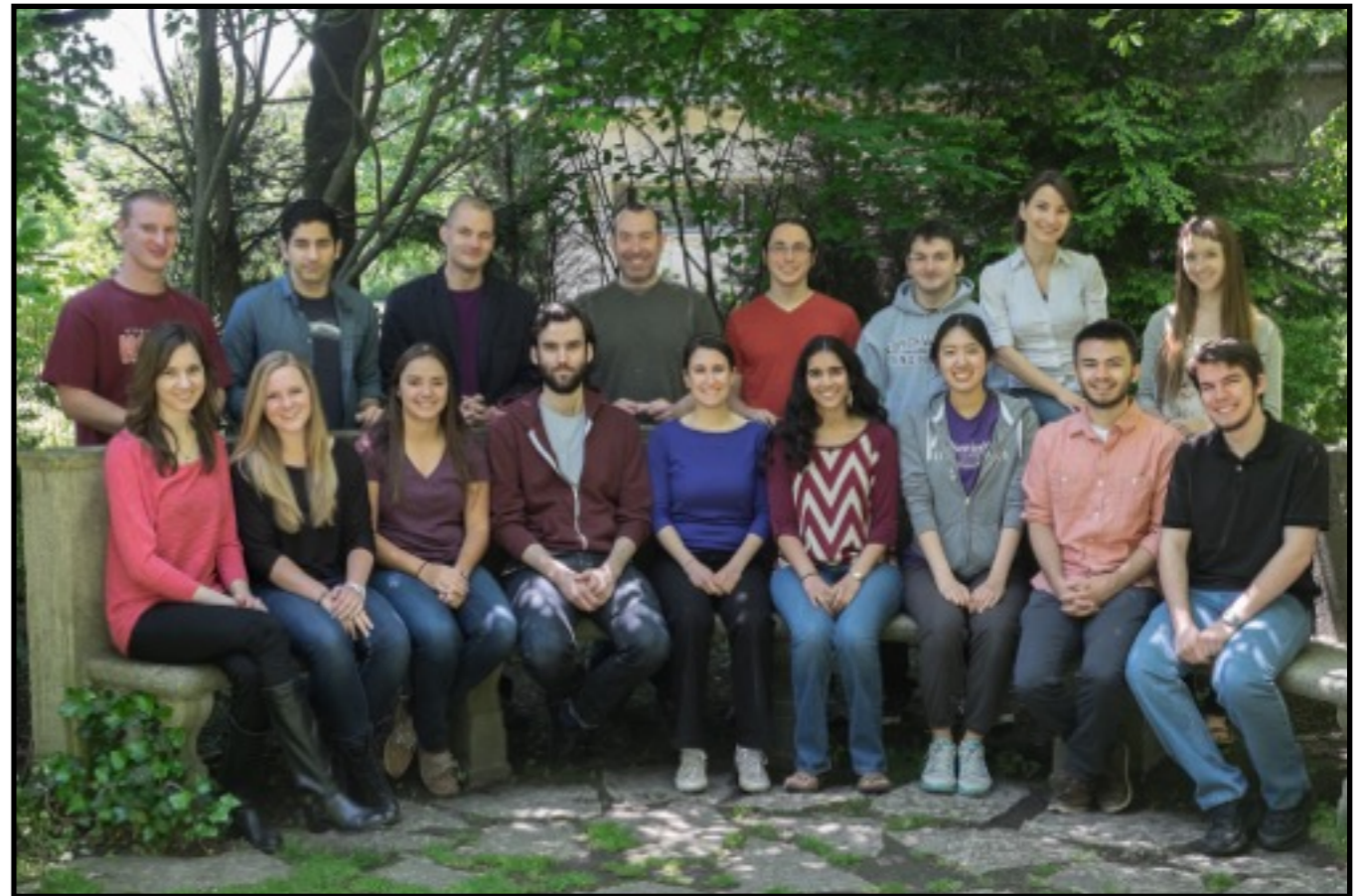




Acknowledgments

Andersen Lab

- Erik Andersen
- Robyn Tanny
- And the rest!



Rockman Lab (NYU)

- Matt Rockman
- Luke Noble
- David Riccardi



Million Mutation Project

