PCAWG-6: Structural Variations

Peter Campbell
Rameen Beroukhim
15 March 2015

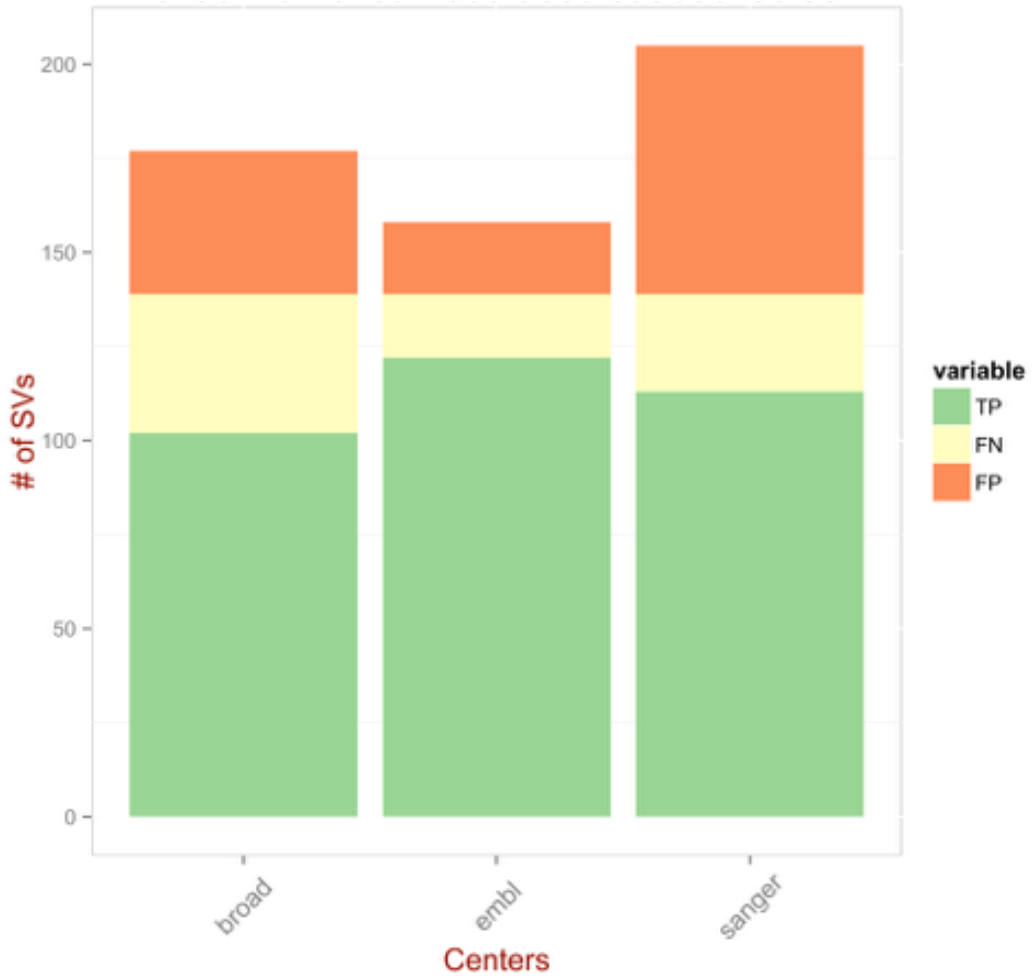
Mission: Describe the landscape of structural variations in cancer

- Determine rearrangement and copy-number alteration profiles across all cancers
- Reconstruct the events that contributed to them
- Detect mechanistic signatures
- Detect signatures of selection
- Determine networks of associated events

Expected outputs

- Lists of consensus rearrangement, telomere length/composition, and retrotransposon calls
- Mechanistic assessments across samples and cancer types
- Lists of rearrangements we think undergo positive selection

Determining structural rearrangements: Integration across different callers



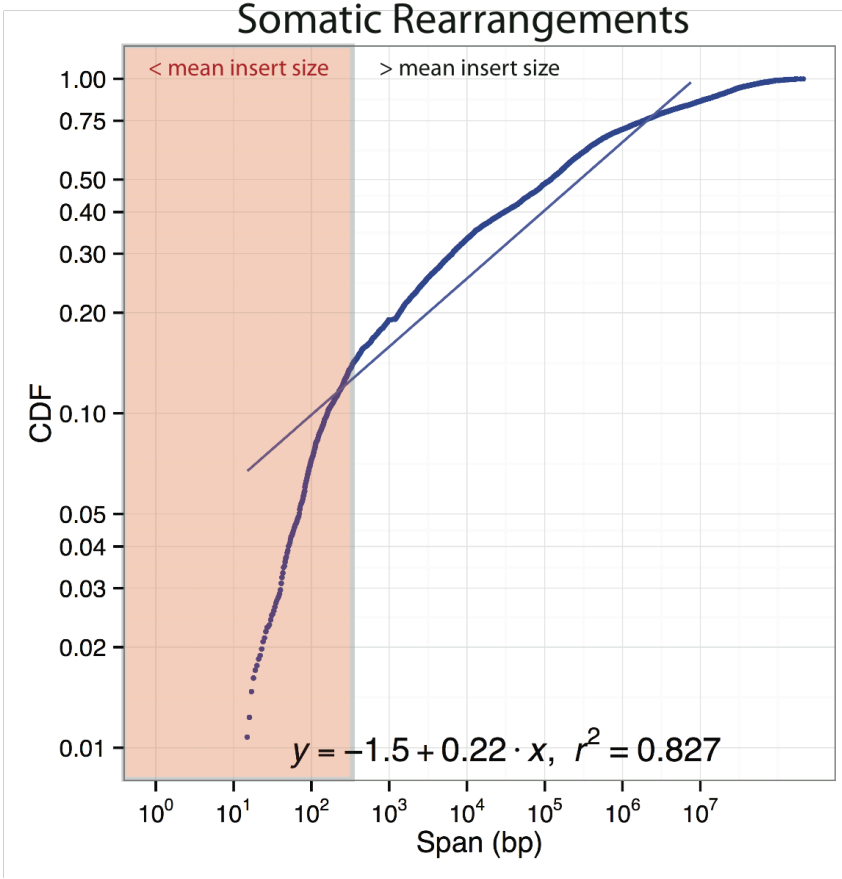
Joachim Weischenfeldt

Size distribution of concordant and discordant calls

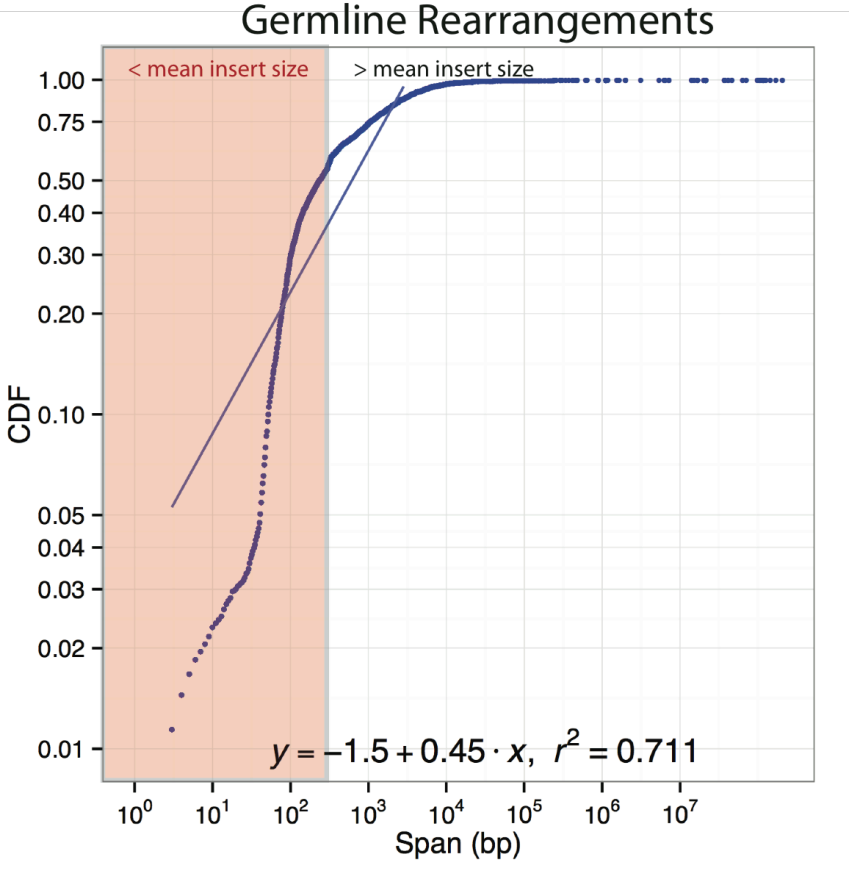
We need to coordinate these results with indel and germline event calling from other groups

Joachim Weischenfeldt

Size distributions: somatic vs germline



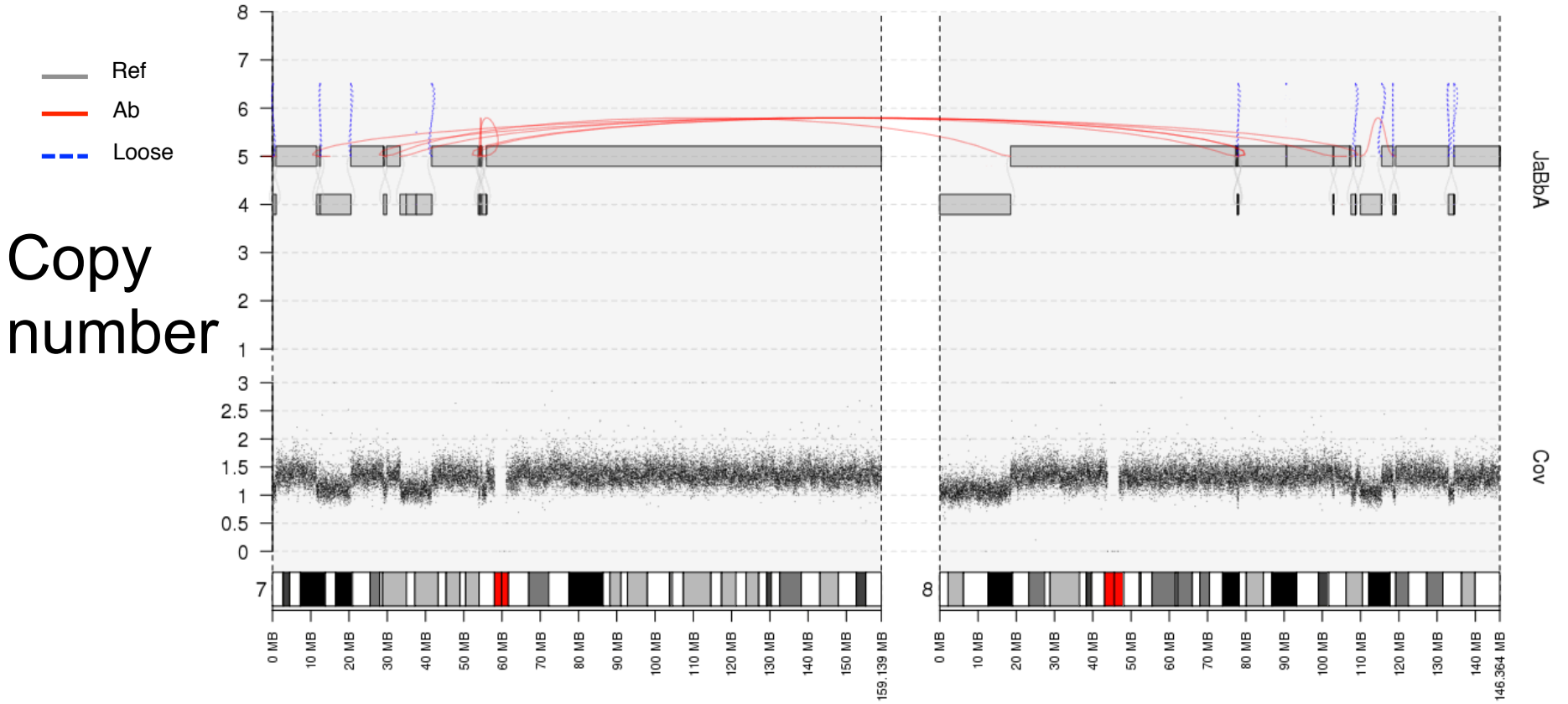
11,483 calls (combined)



51,345 calls (SnowmanSV only)

Jeremiah Wala

Integrating fusions and copy-number states



Marcin Imielinski

Event classification

Events with **zero** breakpoint pairs

- Chromosome gain or loss
- Isochromosome
- Terminal deletion

Events with **2–5** breakpoint pairs

- Inversion
- Insertion
- Retrotransposition
- 'Replication hopping' (inverted-duplication like events)
- Balanced translocation
- Overlaps of simple events

Events with **one** breakpoint pair

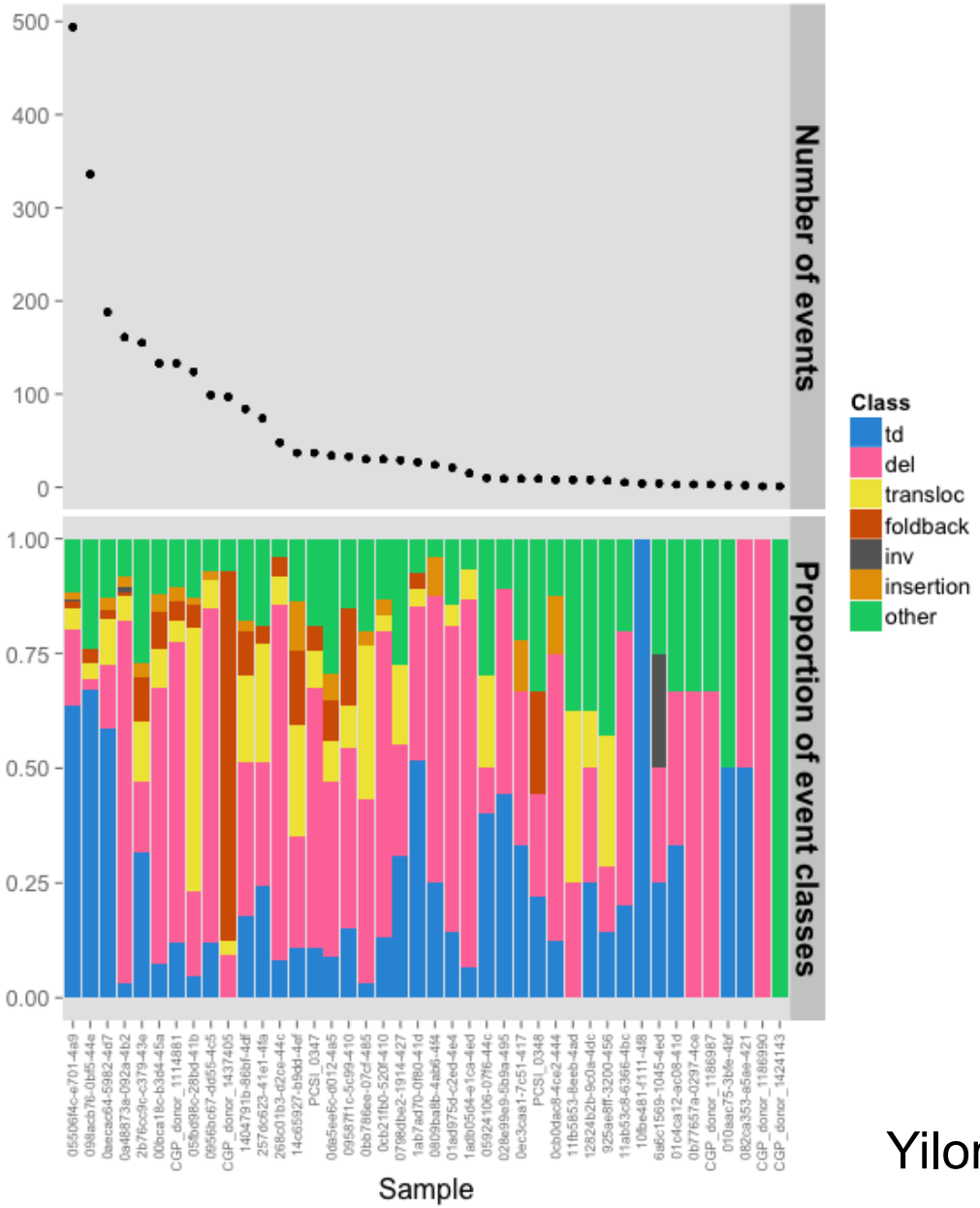
- Tandem duplication
- Interstitial deletion
- Unbalanced translocation
- Foldback (one round of BFB?)
- Double minutes (?)

Events with **many** breakpoint pairs

- Chromothripsis
- Chromoplexy
- Chromoanasynthesis
- BFB cycles
- Events TBD

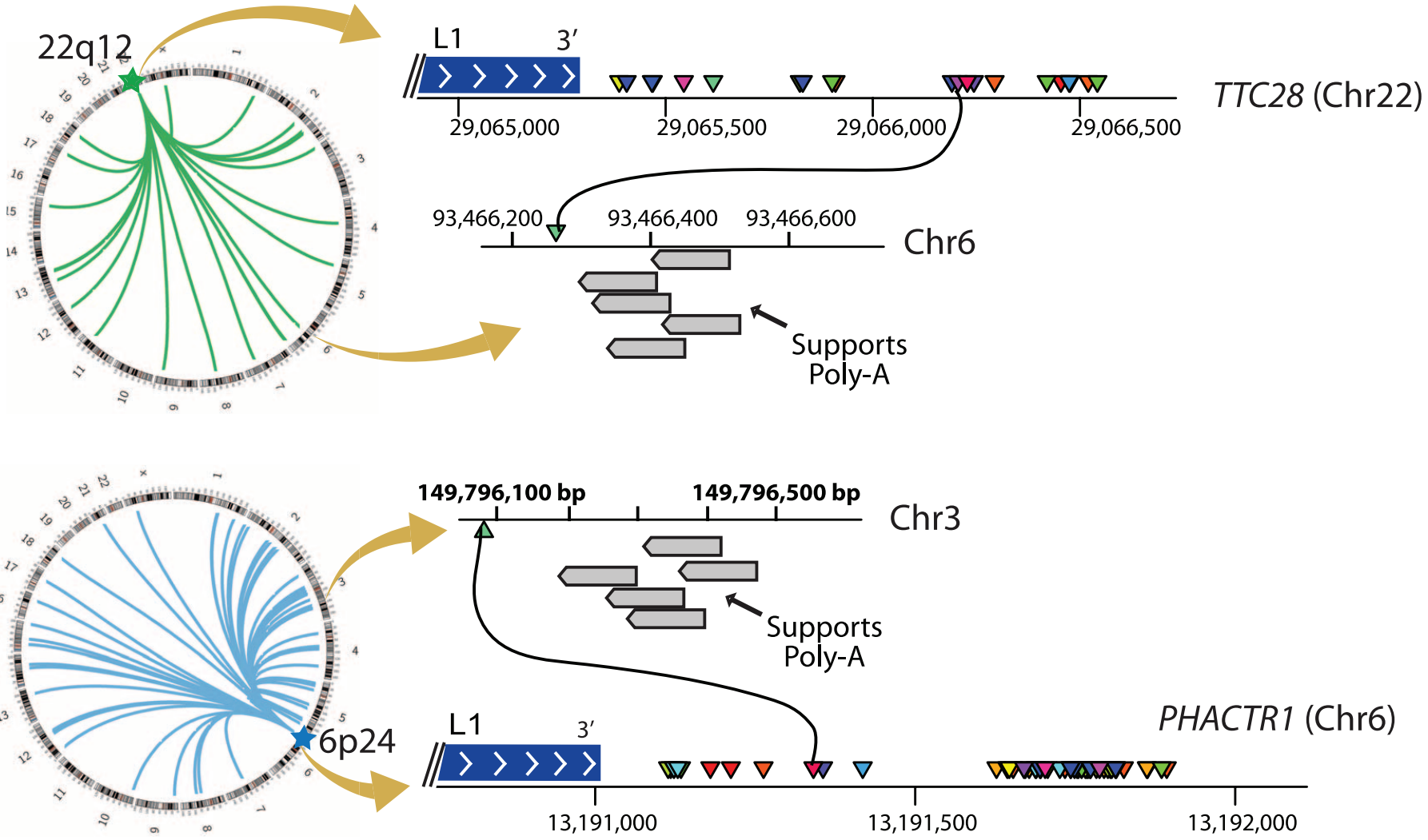
Nicola Roberts

Event calls



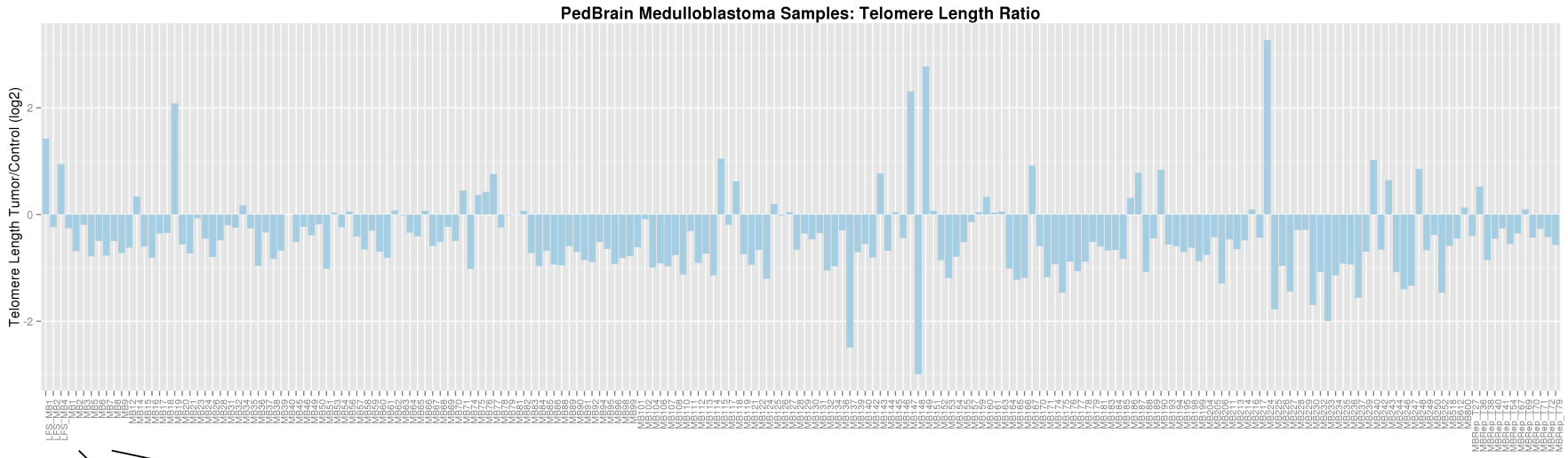
Yilong Yi, Nicola Roberts

L1 retrotransposon-associated rearrangements

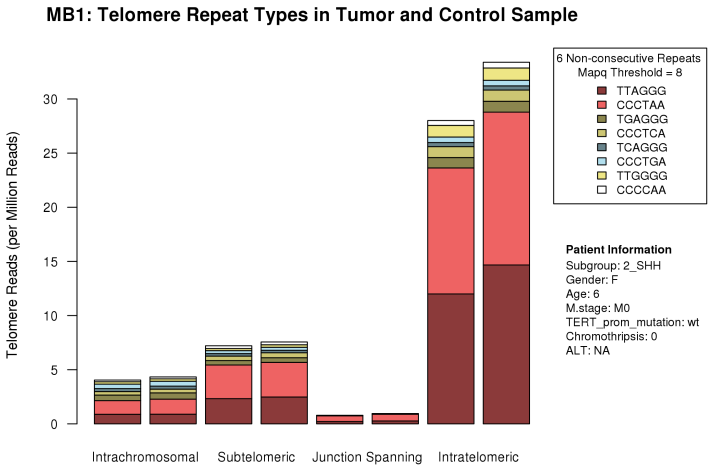


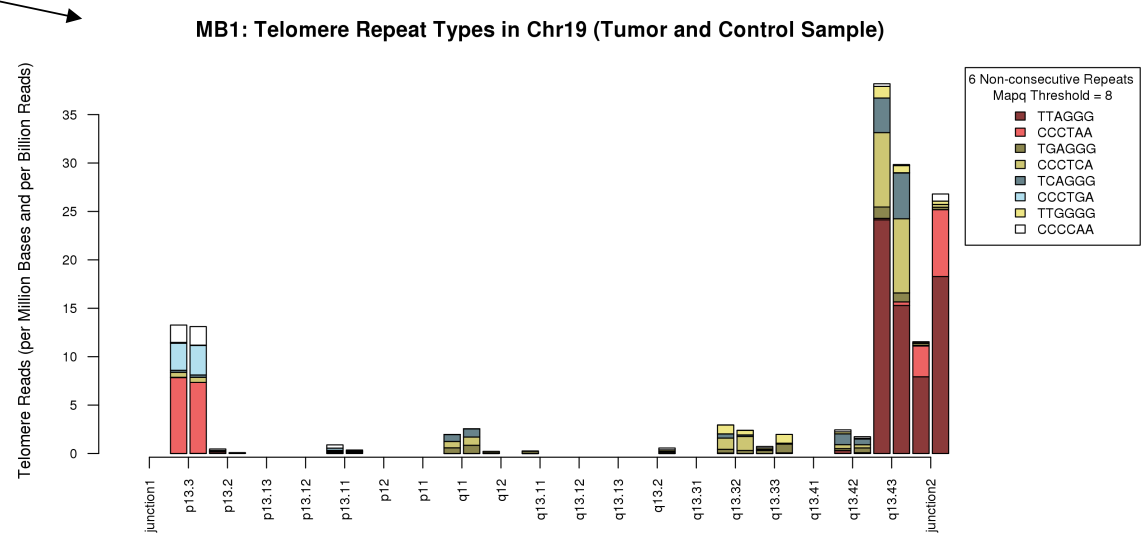
Tubio et al, *Science* 2014

Telomere length and composition



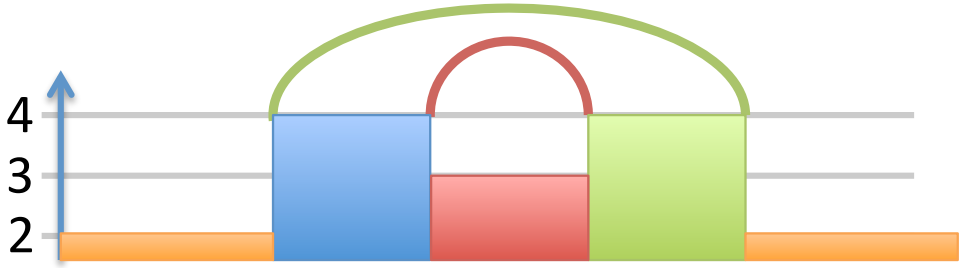
Cohort-wise and patient-wise reports





Lars Feuerbach

Determining the sequence of events



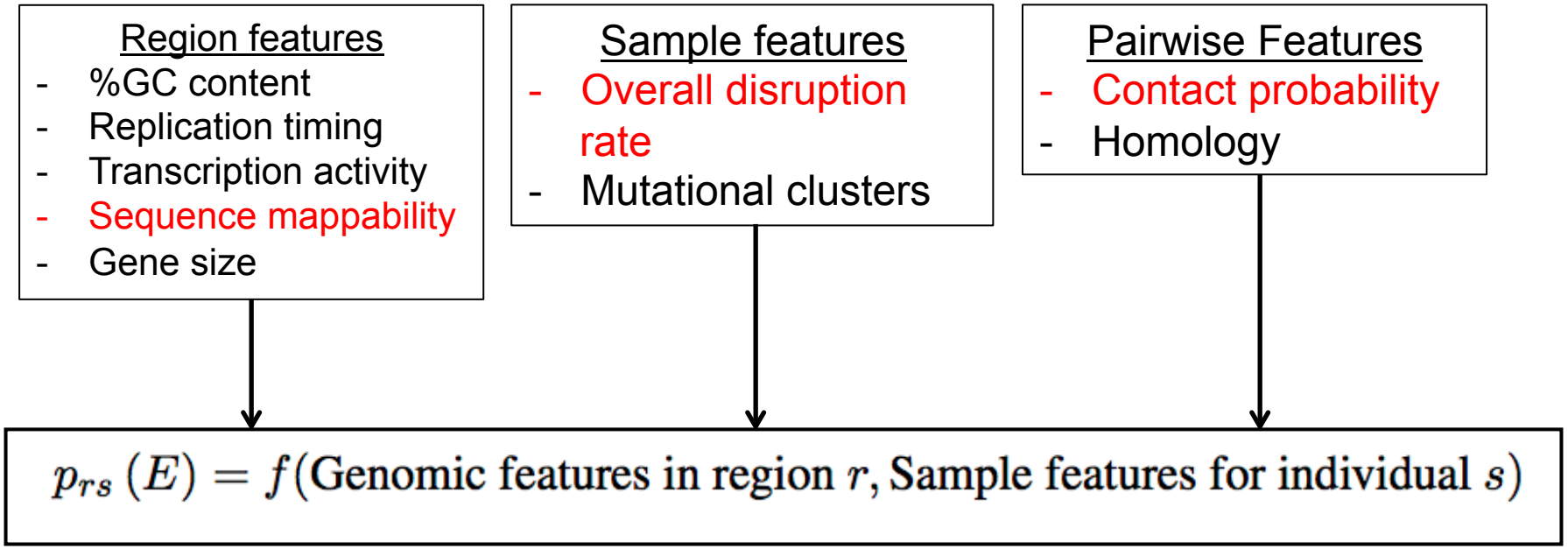
2-4-3-4-2 2-,4+/2+,2-

Parsimonious solution: 2xTD

Evolutions with 3 rearrangements:
TD-TD-del
TD-WCG-UT

Yilong Li

Signatures of selection: developing an initial background model



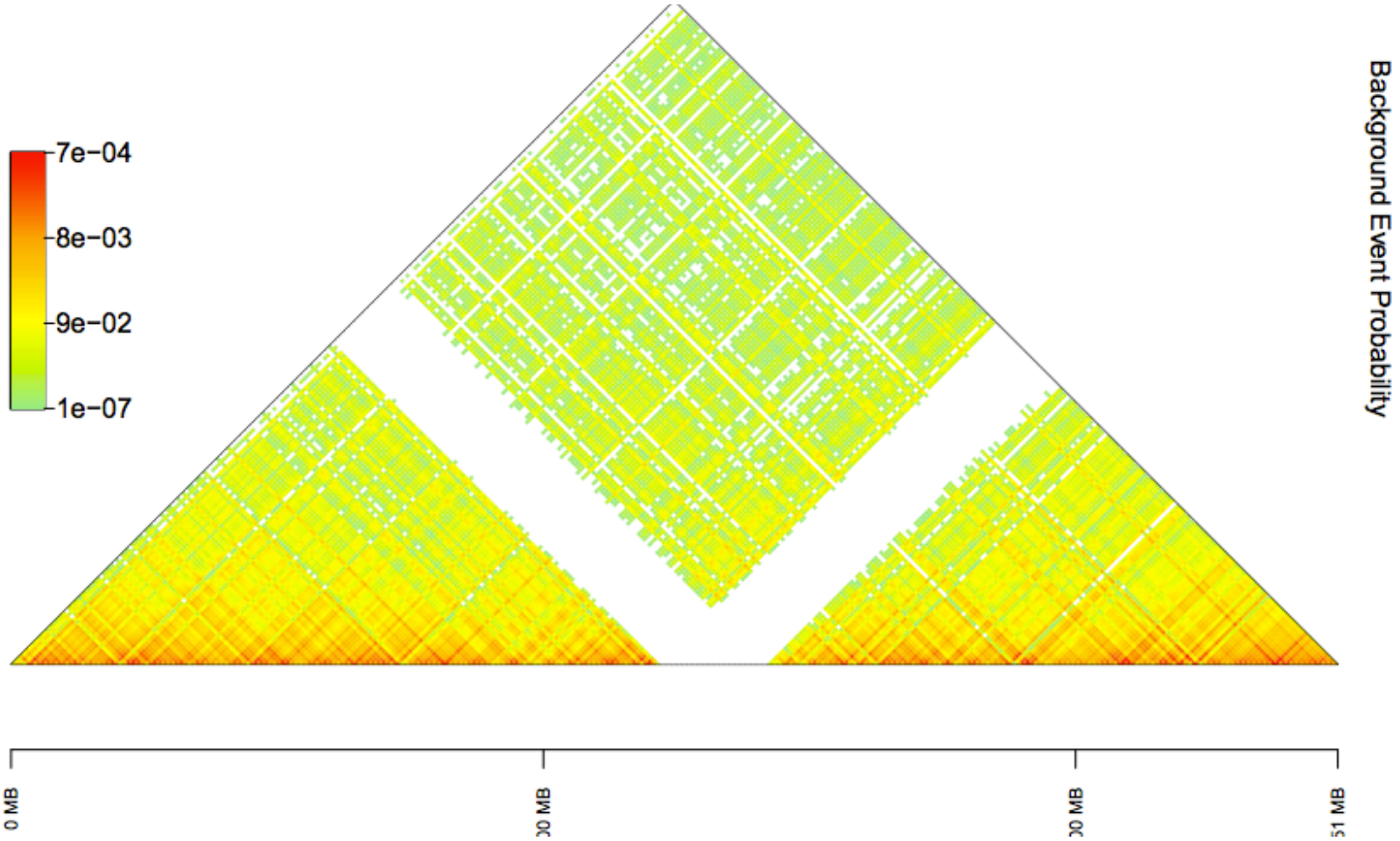
Sequence mappability: Probability proportional to mappable area (100-mer alignability)

Overall disruption rate: Probability proportional to disruption rate

Contact probability: Probability proportional to 1/Length

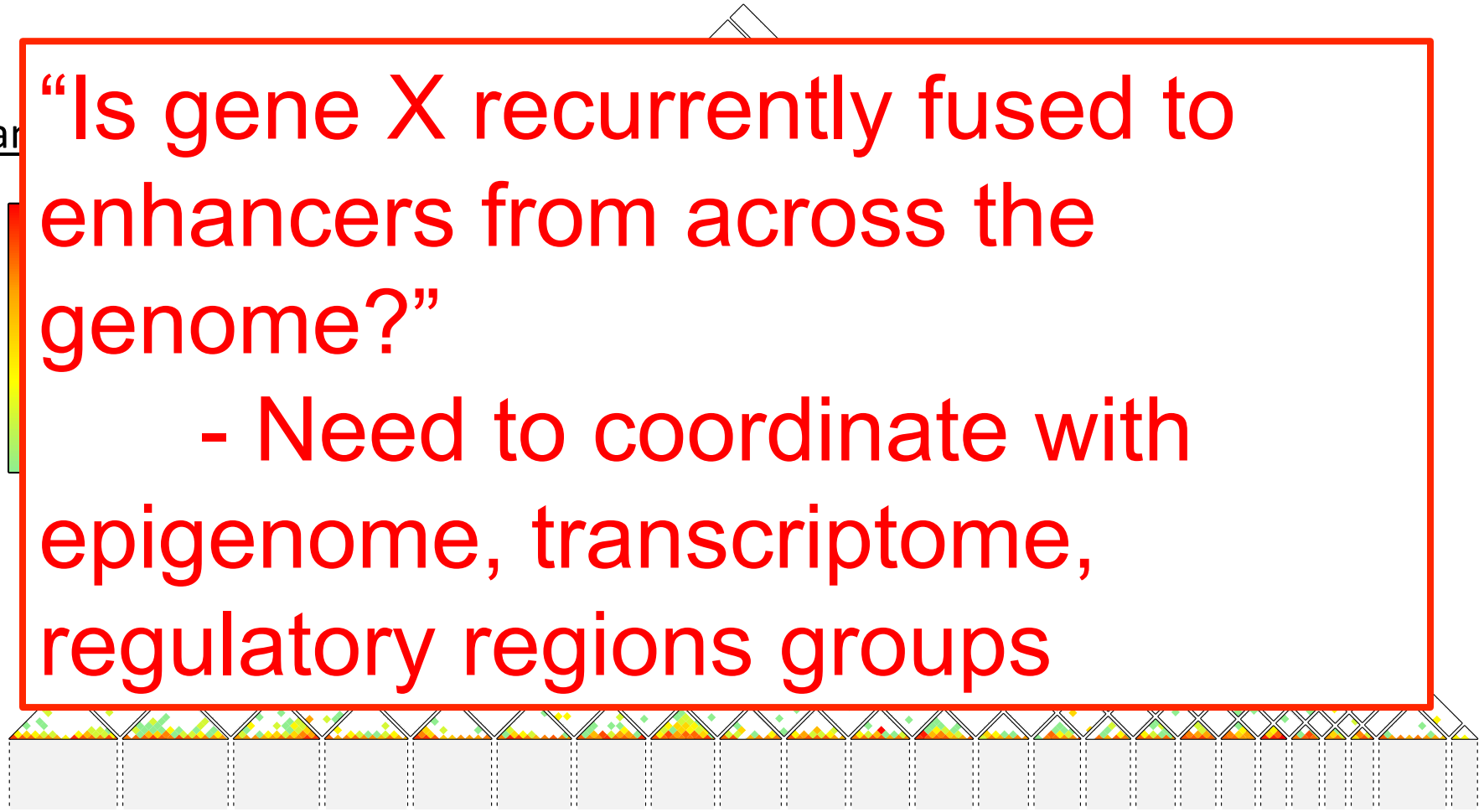
Jeremiah Wala

Chr1: Background event probability
(1/L contact probability model + 100mer alignability)



Jeremiah Wala

Detecting signatures of selection: recurrent fusion pairs

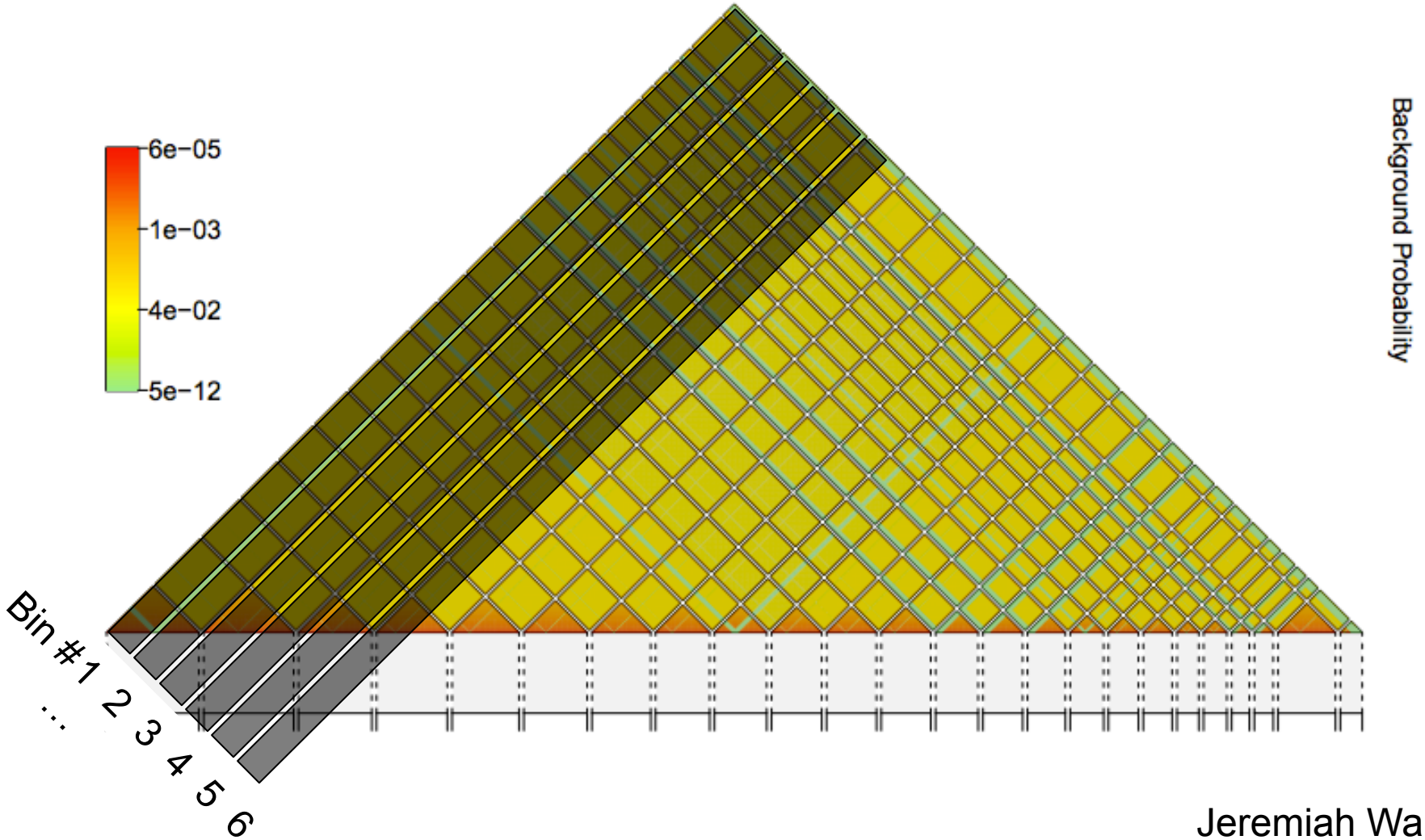


"Is gene X recurrently fused to enhancers from across the genome?"

- Need to coordinate with epigenome, transcriptome, regulatory regions groups

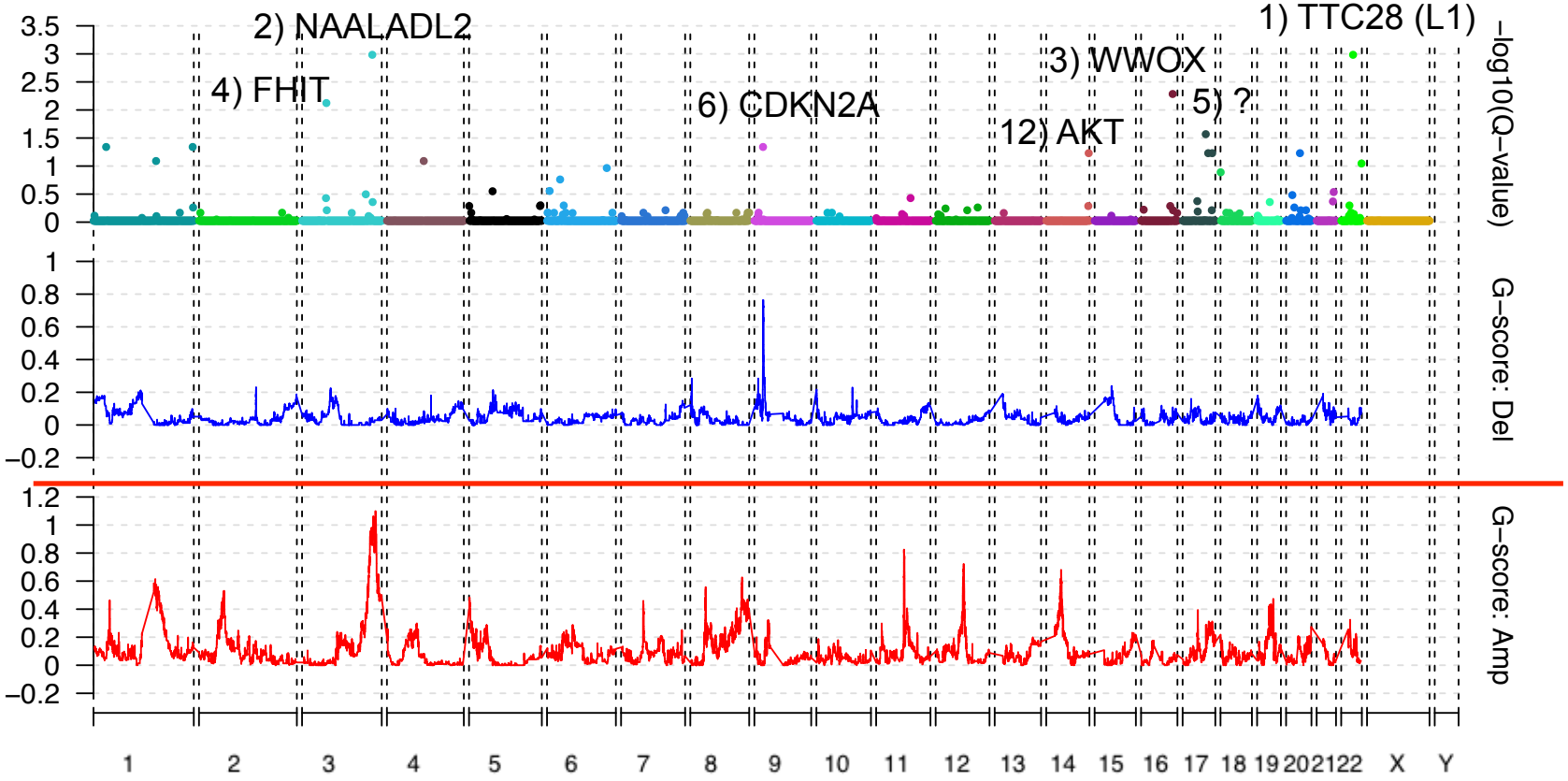
Jeremiah Wala

Probability per 1D bin: Sum across 2D



Jeremiah Wala

Detecting signatures of selection: rearrangements vs copy number

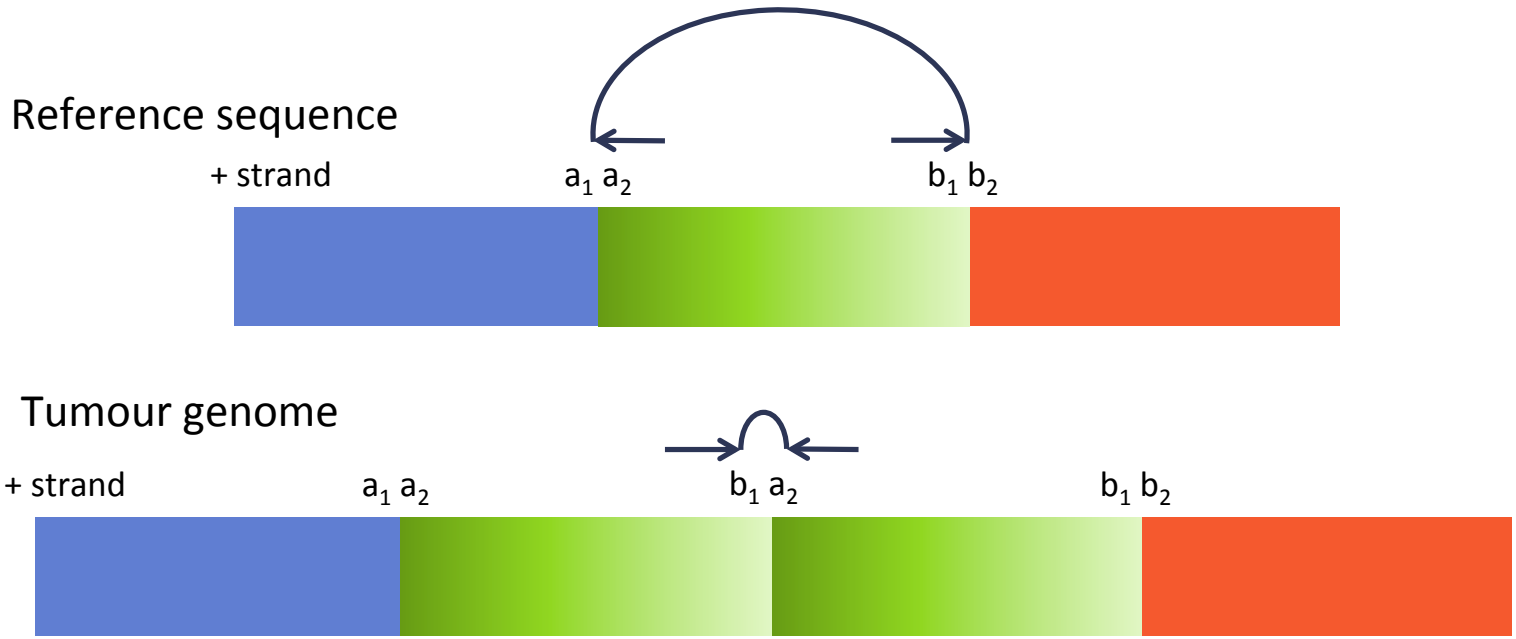


Jeremiah Wala

Classifying by alteration's effects

Example: tandem duplication with breakpoint pair positions a and b:

- Copy number gain of features wholly within a–b
- Disruption of any feature crossing a or b
- Possible fusion gene across $b_1 \wedge a_2$
- Apposition of feature pairs that fall $(< b, > a)$
- Separation of feature pairs that fall $(< a, > b)$

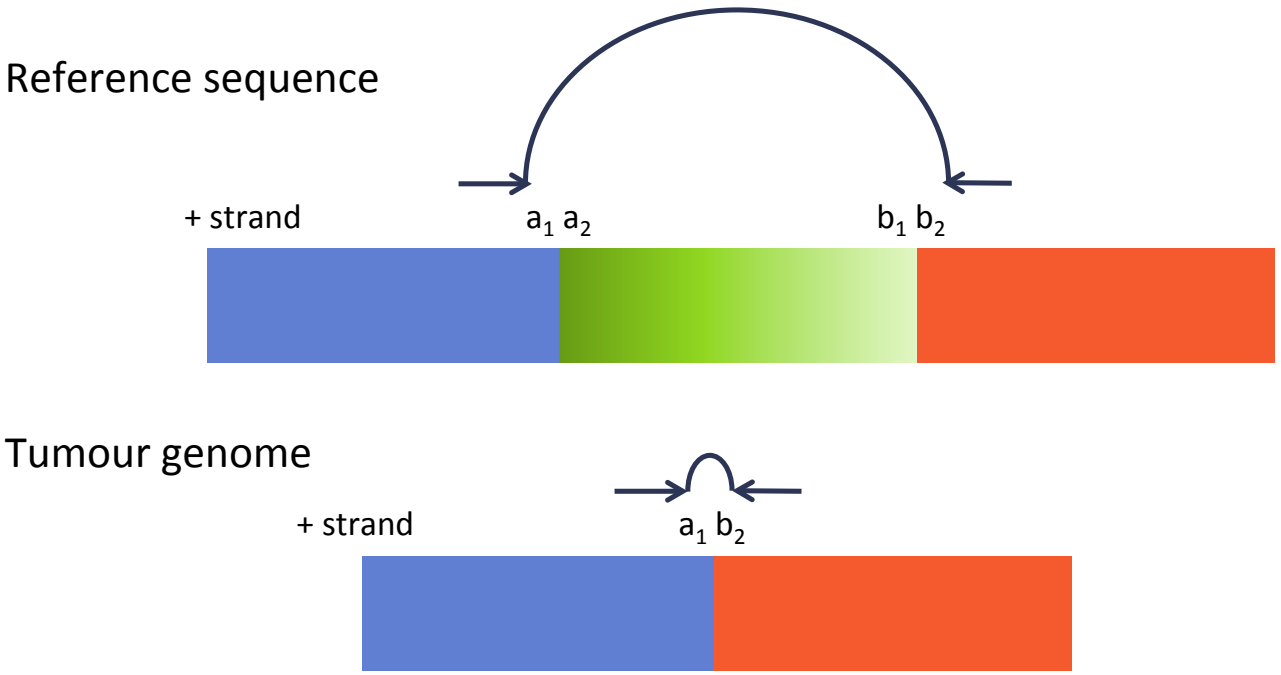


Nicola Roberts

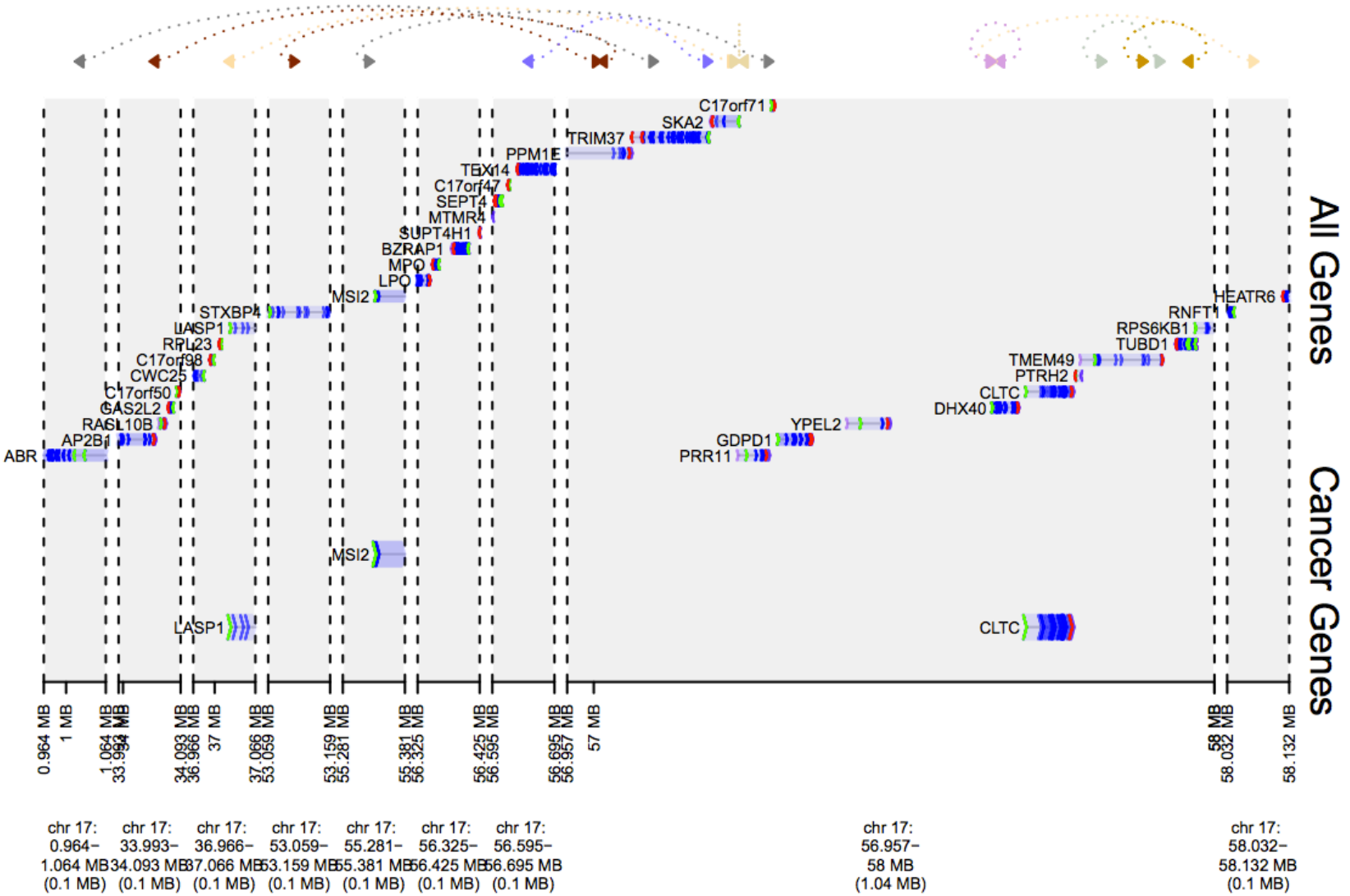
Classifying by alteration's effects

Example: interstitial deletion with breakpoint pair positions a and b:

- Copy number loss of features wholly within a–b
- Disruption of any feature crossing a or b
- Possible fusion gene across $a_1 \wedge b_2$ (allowing for possible nts or mh)
- Apposition of feature pairs that fall $(< a, > b)$
- Separation of feature pairs that fall $(< a, > a)$ and $(< b, > b)$



Common features of rearrangements in hot regions?



Jeremiah Wala