# Estimation of genomic inbreeding in cattle: the impact of SNP chip density

*Maja Ferencakovic*, Ino Curik

and Johann Sölkner

The 63rd Annual Meeting of the European Association for Animal Production
August 27th - 31st, Bratislava, Slovakia

# Estimation of inbreeding

- **Traditionally - using pedigree records – $F_{PED}$**

  - Disadvantages

- **Use of dense marker to estimate F**

  - Several advantages over $F_{PED}$

  - Excess of SNP homozygosity, Van Raden 2008, Yang *et al.* 2010

  - Runs of homozygosity - $F_{ROH}$

# Runs of homozygosity (ROH)

- **Long stretches of homozygous SNP**

# Runs of homozygosity (ROH)

- **Long stretches of homozygous SNP**
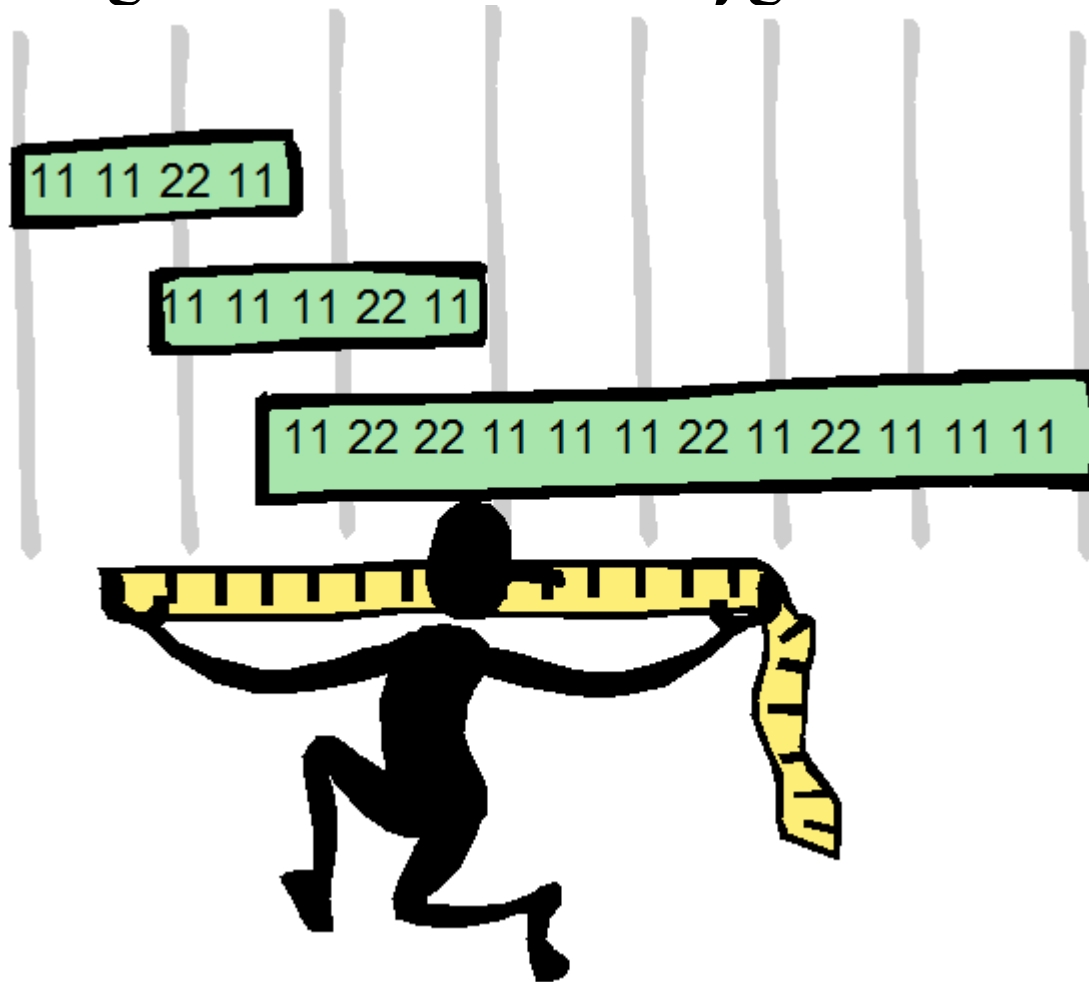
# Runs of homozygosity (ROH)

- **Long stretches of homozygous SNP**
  - Present because of transmission of identical haplotypes from parents to offspring
  - Their length is affected by recombination events
    - Segment Length (M) = 1/2*g $\longrightarrow$ g = 1/2*Segment Length (M)
  - Are used as measure of homozygosity (autozygosity)

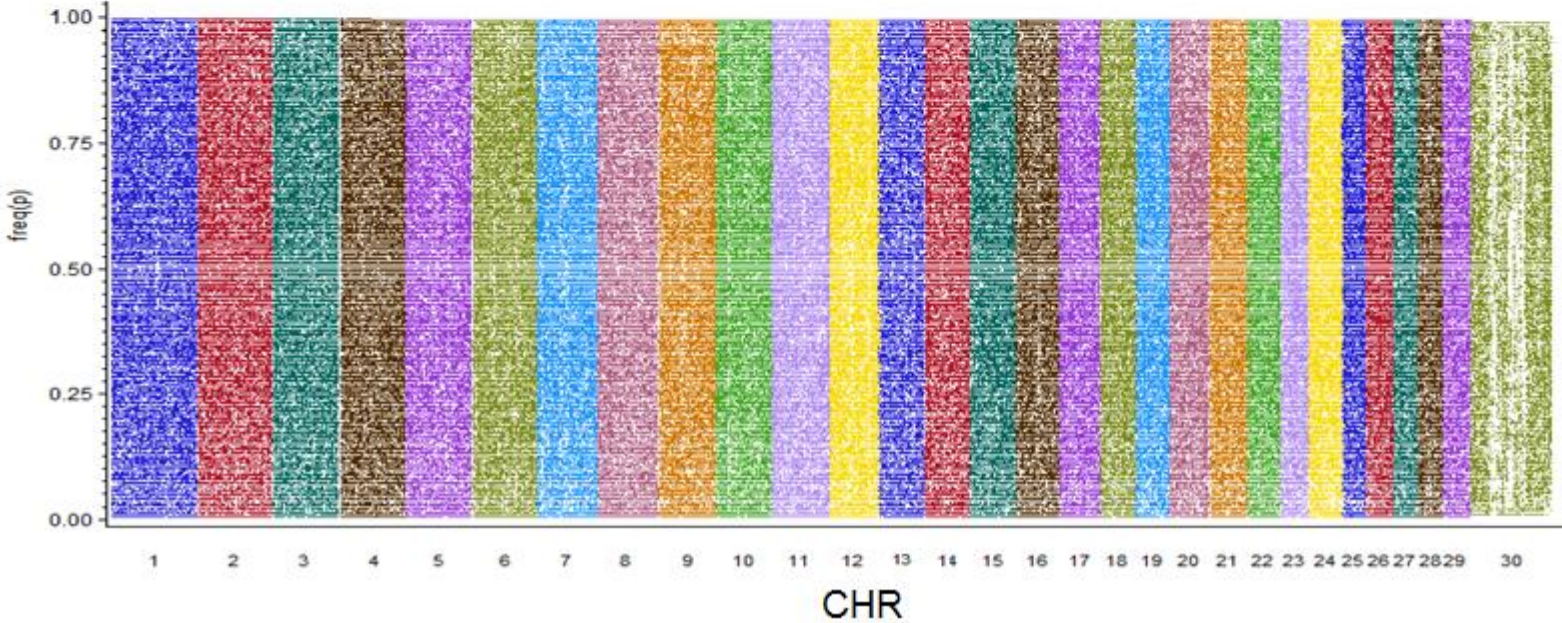$$F_{ROHk} : \frac{\sum_k length\ (ROH_k)}{L}$$

# SNP chip density

- **Most used :**

  - Illumina BovineSNP50 Genotyping BeadChip (54k panel)

  - BovineHD Genotyping BeadChip (HD panel)

  - 45 $ vs. 150 $

  - Difference?

# SNP chip density
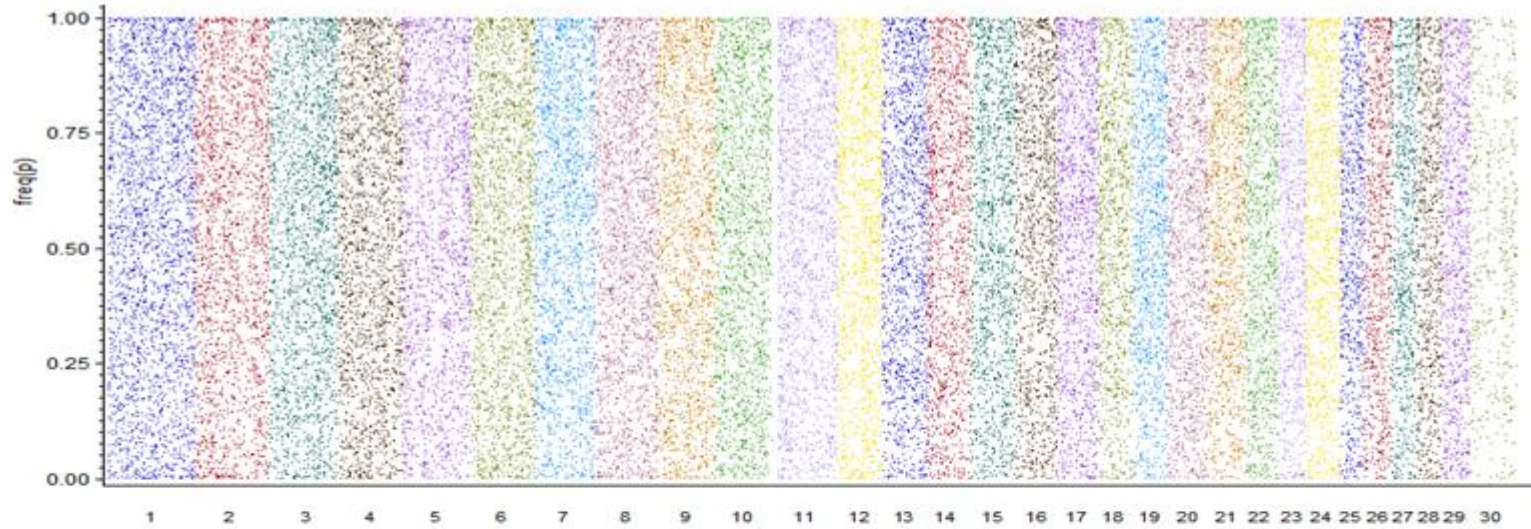
# SNP chip density

- **Most used :**

  - Illumina BovineSNP50 Genotyping BeadChip (54k panel)

  - BovineHD Genotyping BeadChip (HD panel)

  - 45 $ vs. 150 $

  - Difference?

- **Questions :**

  - Impact of SNP chip density on ROH estimation

  - Is 54k chip appropriate for identifying ROH?

# Genotype data

- **117 Tyrolean Grey bulls genotyped with 54k and HD panel**

- **We excluded using SNP & Variation Suite v7.6.8**
  (Golden Helix, Bozeman, MT, www.goldenhelix.com ):

  - SNP with gc_score $\leq 0.7$

  - SNP with cluster separation score $\leq 0.4$

  - SNP unassigned to chromosome or without bp position

  - SNP from X chromosome

  - SNP with > 25% of missing genotypes

  - Animals with > 5% of missing genotypes

GOLDEN HELIX

# Genotype data

- **115 animals**

- **Final 54k panel data set:**

  - 41775 autosomal SNP

  - Autosomal SNP covering 2.51 Gb of genome

- **Final HD panel data set:**

  - 684172 autosomal SNP

  - Autosomal SNP covering 2.52 Gb of genome

# Definition of a run of homozygosity

- **Defined with SNP & Variation Suite v7.6.8**

- **To call a ROH**: min. 15 homozygous SNP in row

    : min. required density  1 SNP every 100kb

    : max. gap between 2 SNP ≤ 1000kb

- **5 ROH length categories were used:**

| Parameter | 1-2 Mb | 2-4 Mb | 4-8 Mb | 8-16 Mb | > 16 Mb |
|---|---|---|---|---|---|
| # missing allowed HD/54k | 4/0 | 8/0 | 16/1 | 32/2 | 64/4 |
| # heterozygotes allowed HD/54k | 1/0 | 2/0 | 4/0 | 8/0 | 16/1 |

GOLDEN HELIX

*Accelerating the Quest for Significance™*

# Definition of a run of homozygosity

- **Defined with SNP & Variation Suite v7.6.8**

- **To call a ROH**: min. 15 homozygous SNP in row

  : min. required density  1 SNP every 100kb

  : max. gap between 2 SNP ≤ 1000kb

- **5 ROH length categories were used:**

| Parameter | 1-2 Mb | 2-4 Mb | 4-8 Mb | 8-16 Mb | > 16 Mb |
|---|---|---|---|---|---|
| # missing allowed HD/54k | 4/0 | 8/0 | 16/1 | 32/2 | 64/4 |
| # heterozygotes allowed HD/54k | 1/0 | 2/0 | 4/0 | 8/0 | 16/1 |

# ROH inbreeding coefficients

- **From 5 ROH length categories we calculated 5 inbreeding coefficients:**

  - $F_{ROH > 1\ Mb}$ $\longrightarrow$ up to 50 generations from common ancestor

  - $F_{ROH > 2\ Mb}$ $\longrightarrow$ up to 20 generations

  - $F_{ROH > 4\ Mb}$ $\longrightarrow$ up to 12.5 generations

  - $F_{ROH > 8Mb}$ $\longrightarrow$ up to $\approx 6$ generations

  - $F_{ROH > 16\ Mb}$ $\longrightarrow$ up to $\approx 3$ generations

# Pedigree data

- Pedigree included 1647 animals

- Mean complete generation equivalent $= 6.23$

- $F_{PED}$ mean value $= 0.024 \pm 0.022$ (range $0 - 0.17$)

- Correlations between the $F_{PED}$ and $F_{ROH}$ from

  - 54k panel $= 0.79 - 0.82$ (highest for $F_{ROH>16Mb}$)

  - HD panel $= 0.76 - 0.81$ (highest for $F_{ROH>16Mb}$)
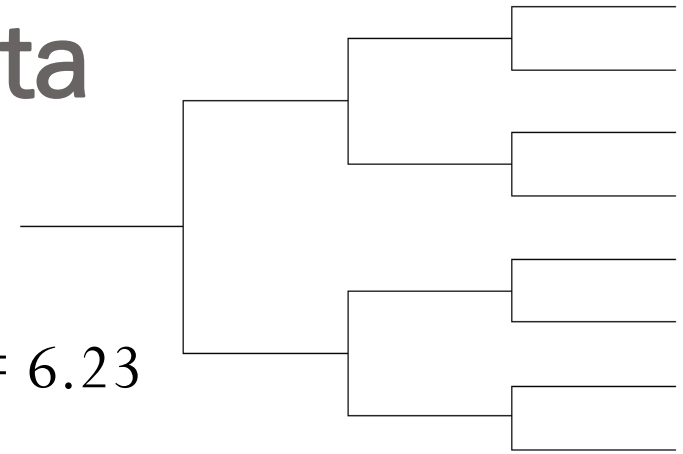
# ROH inbreeding coefficients
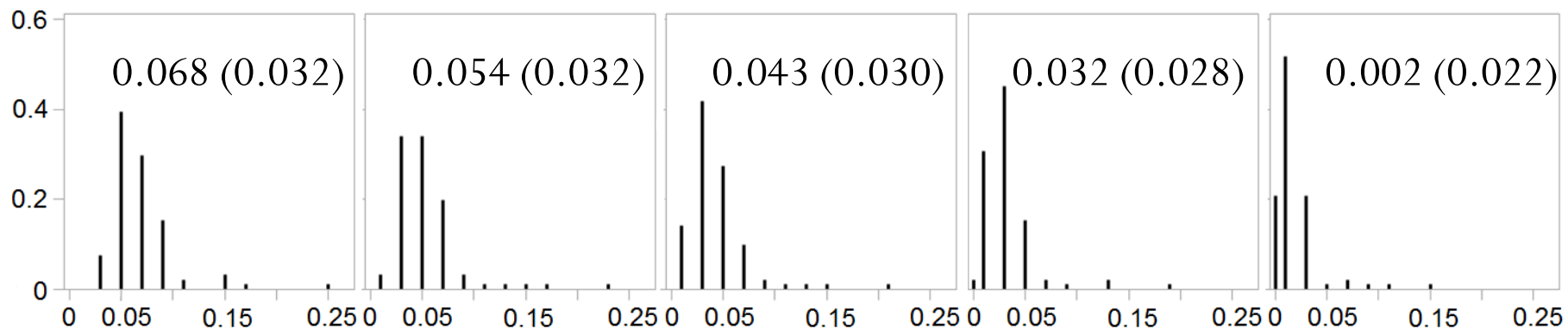


HD panel
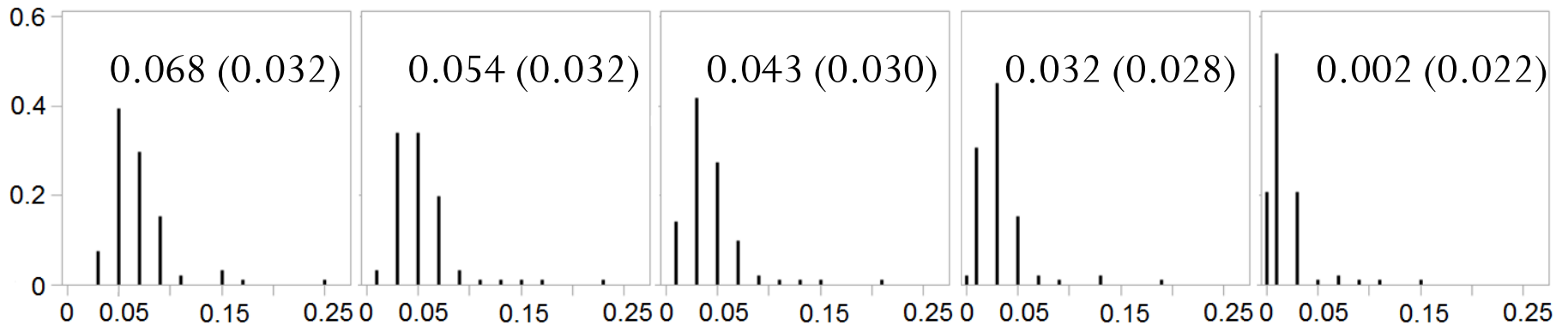
$F_{ROH>1\ Mb}$    $F_{ROH>2\ Mb}$    $F_{ROH>4\ Mb}$    $F_{ROH>8\ Mb}$    $F_{ROH>16\ Mb}$

54k panel

# ROH inbreeding coefficients
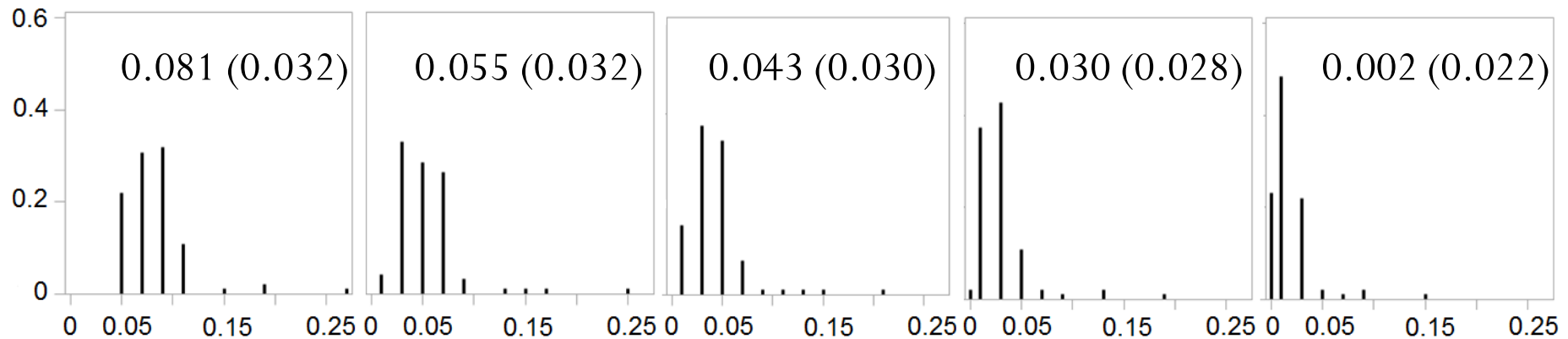
# Small ROH segments – ancient inbreeding



**54k panel**
**ROH >1 Mb**

Chr1:11507.6K  Chr1:34522.8K  Chr1:57537.9K  Chr1:80553.1K  Chr1:103568K  Chr1:126583K  Chr1:149599K

**HD panel**
**ROH >1 Mb**
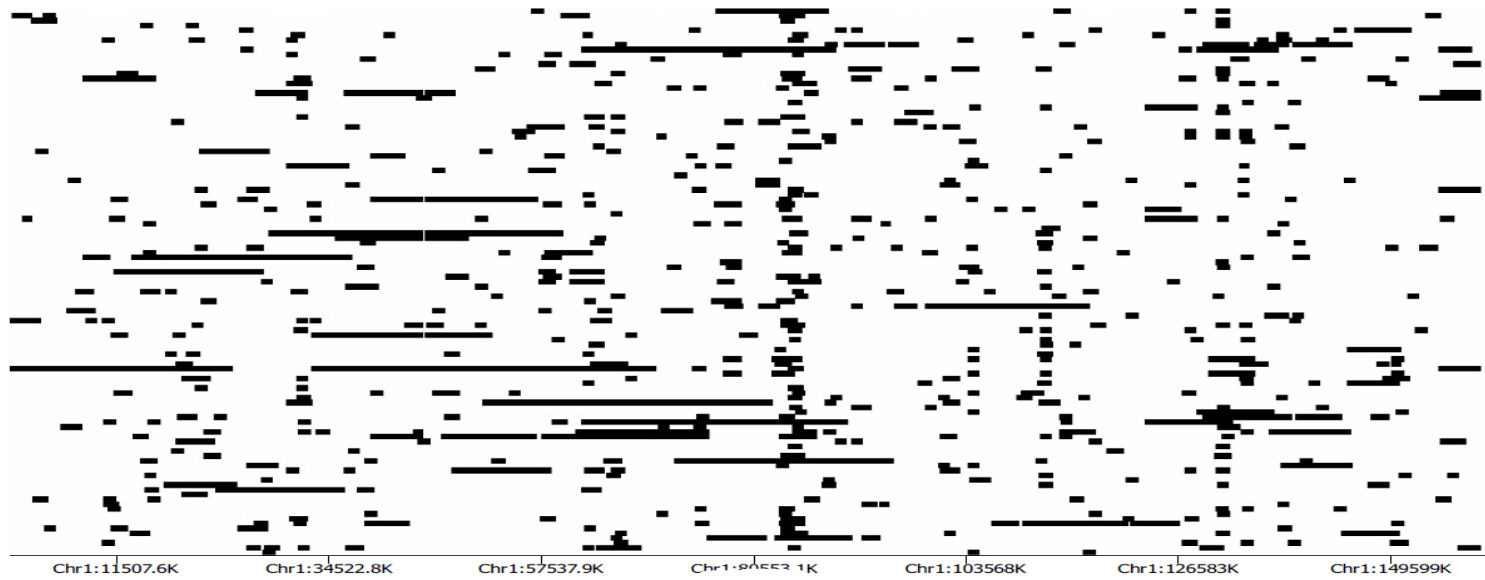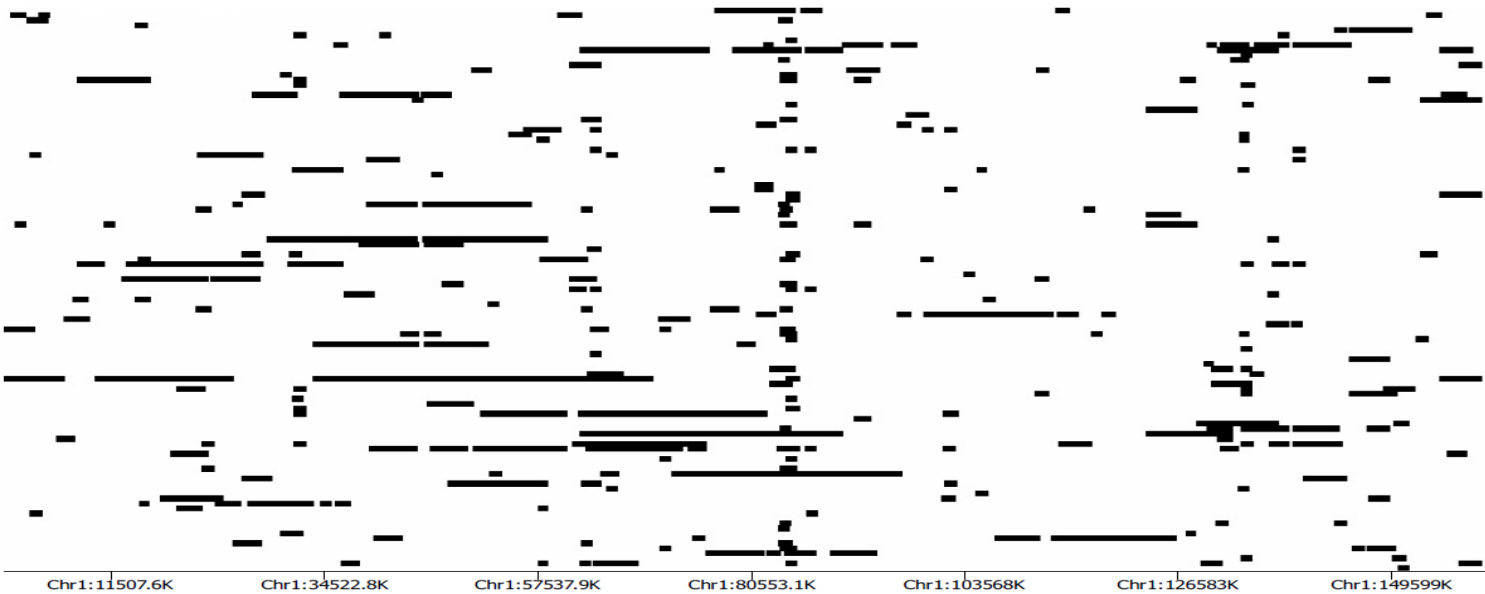
Chr1:11507.6K  Chr1:34522.8K  Chr1:57537.9K  Chr1:80553.1K  Chr1:103568K  Chr1:126583K  Chr1:149599K
Position

# Large ROH segments – recent inbreeding
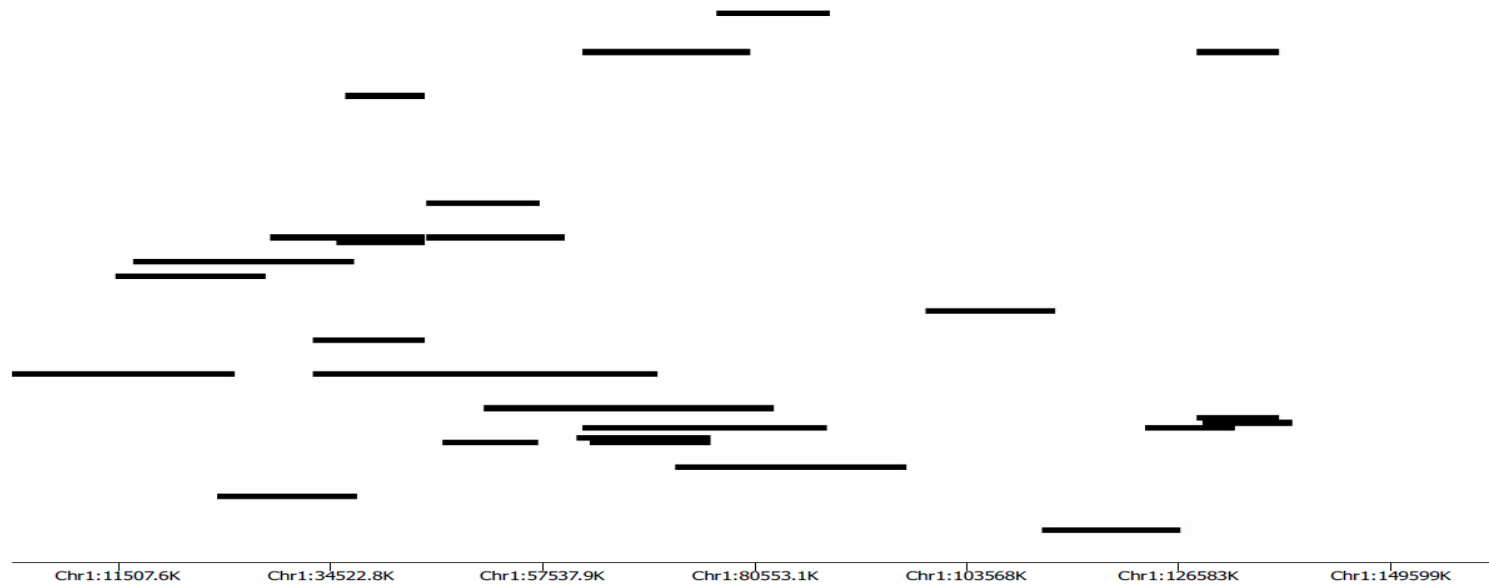


54k panel
ROH >8 Mb

Chr1:11507.6K   Chr1:34522.8K   Chr1:57537.9K   Chr1:80553.1K   Chr1:103568K   Chr1:126583K   Chr1:149599K

HD panel
ROH >8 Mb

Chr1:11507.6K   Chr1:34522.8K   Chr1:57537.9K   Chr1:80553.1K   Chr1:103568K   Chr1:126583K   Chr1:149599K
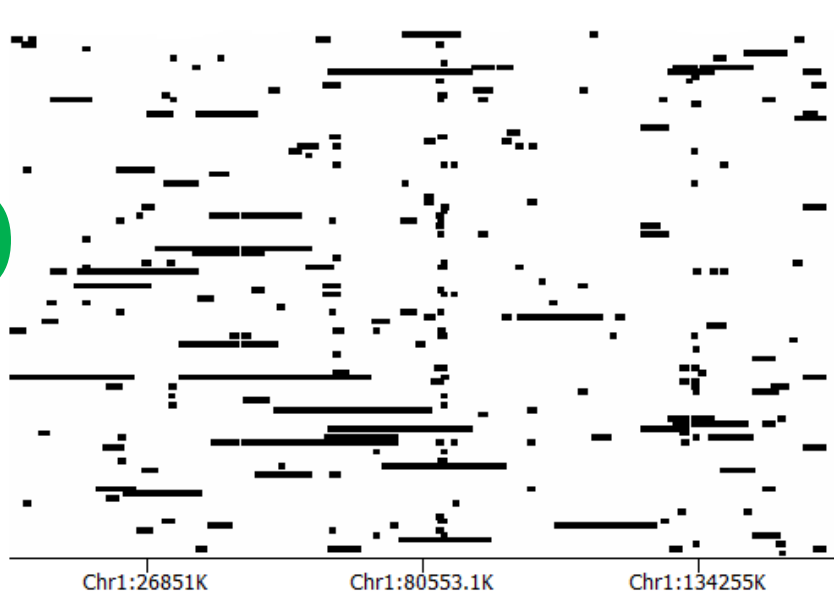Position
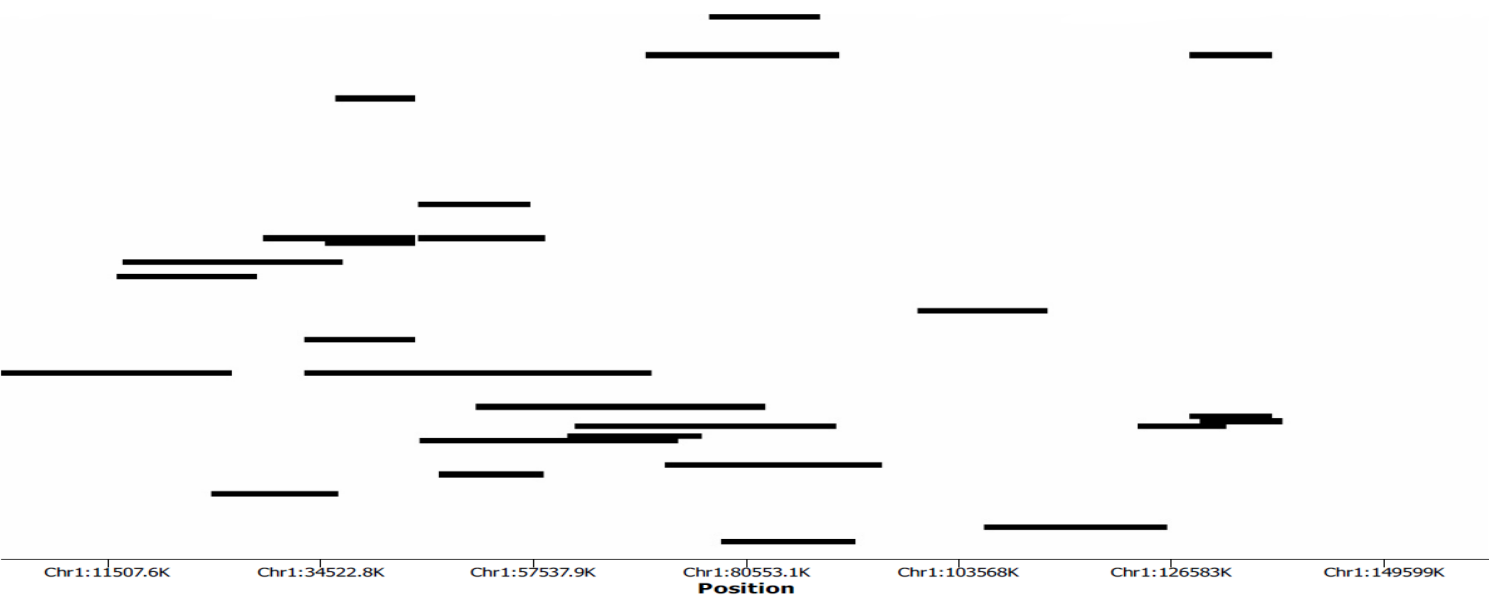
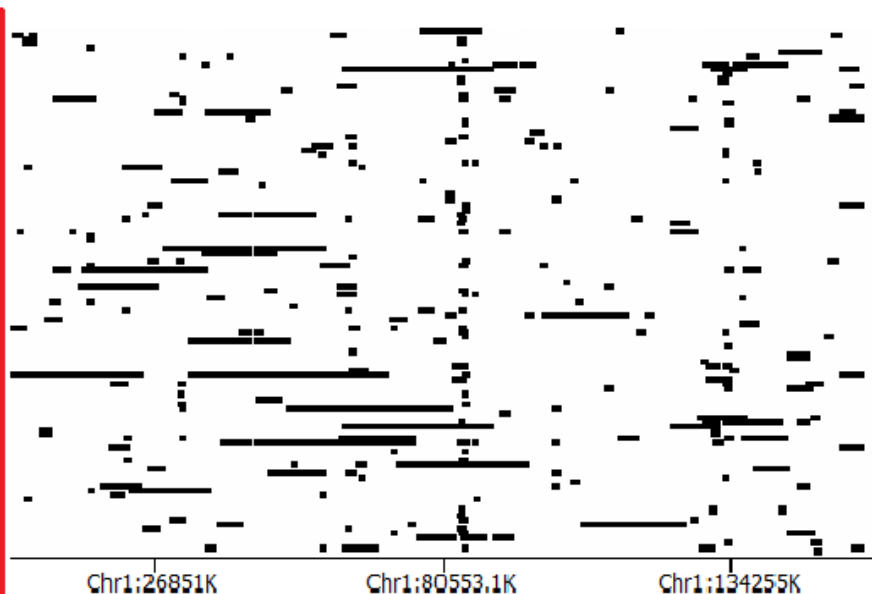# Importance of adjusting parameters

HD panel example for ROH 1-2 Mb

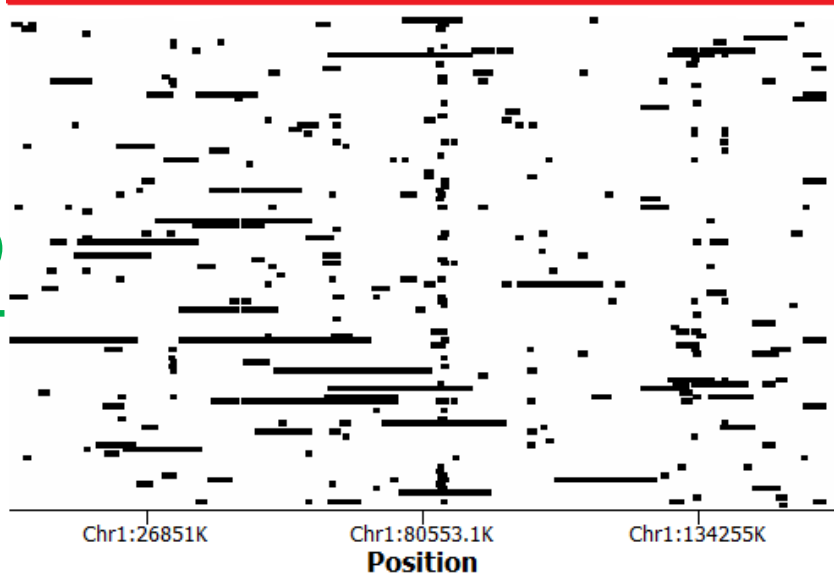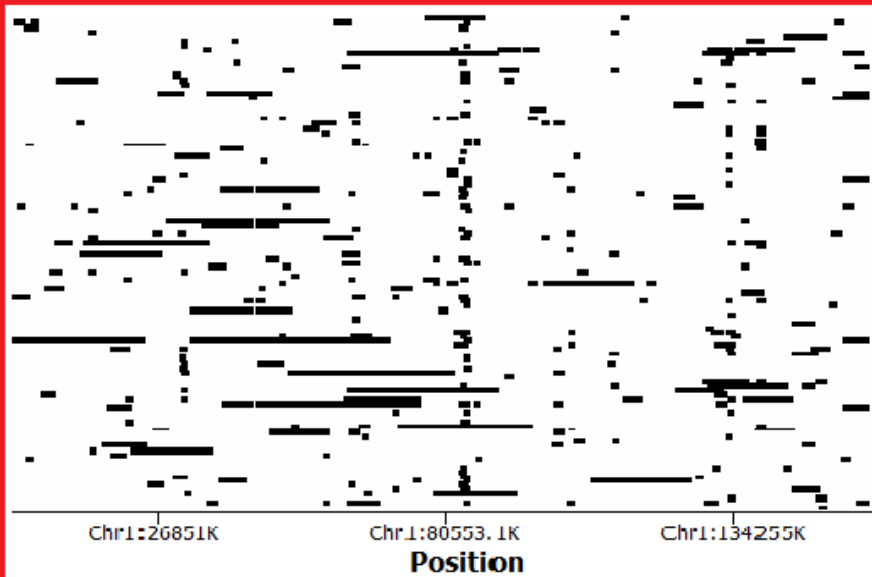Allowing 0, 1, 2 or 4 heterozygotes

# Importance of adjusting parameters

# Importance of adjusting parameters
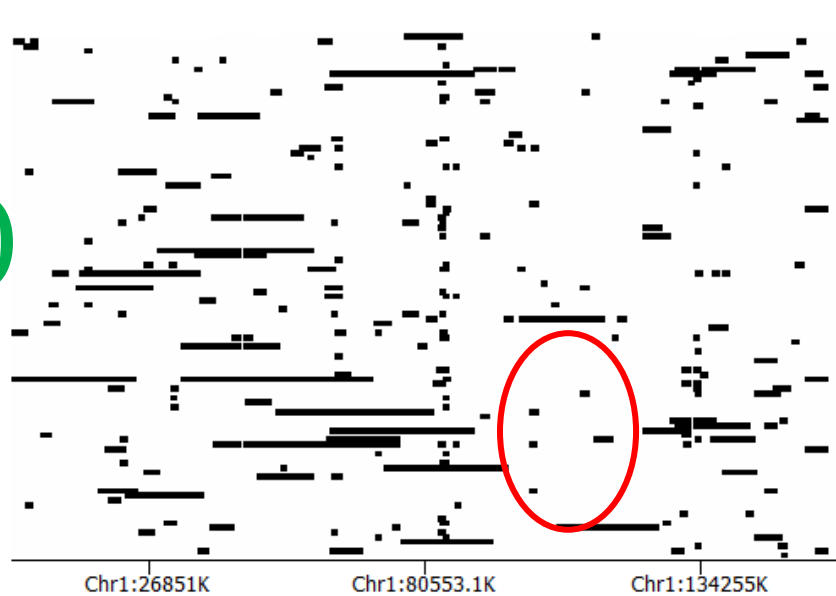
# Importance of adjusting parameters

HD panel example for ROH 8-16 Mb

Allowing 0 or 10 heterozygotes

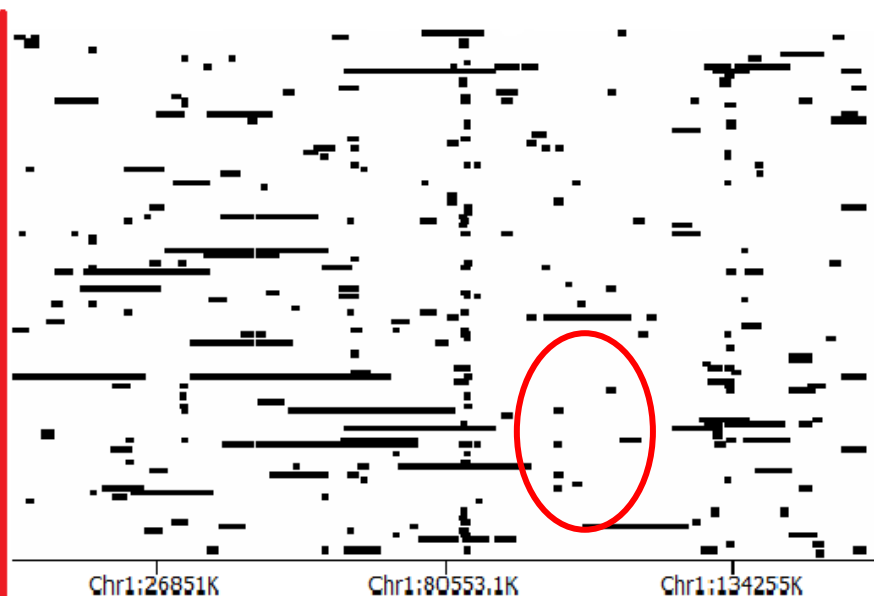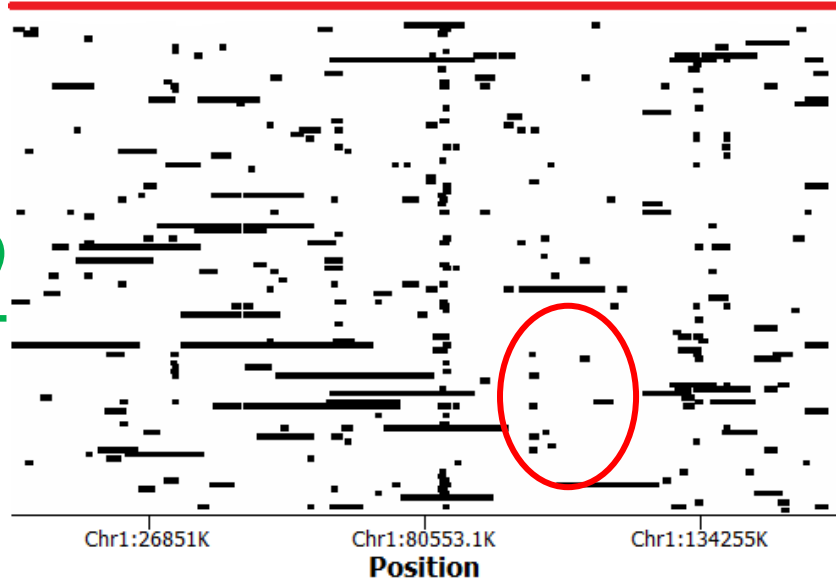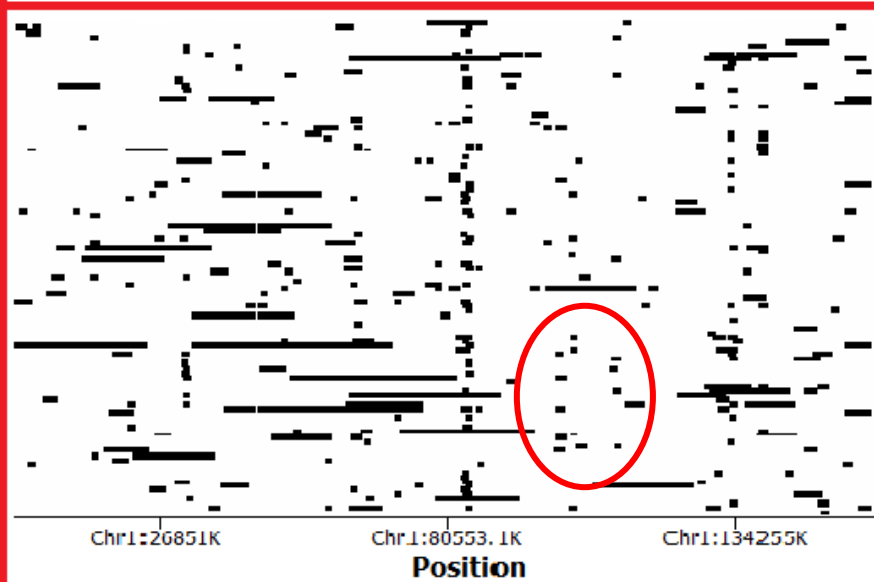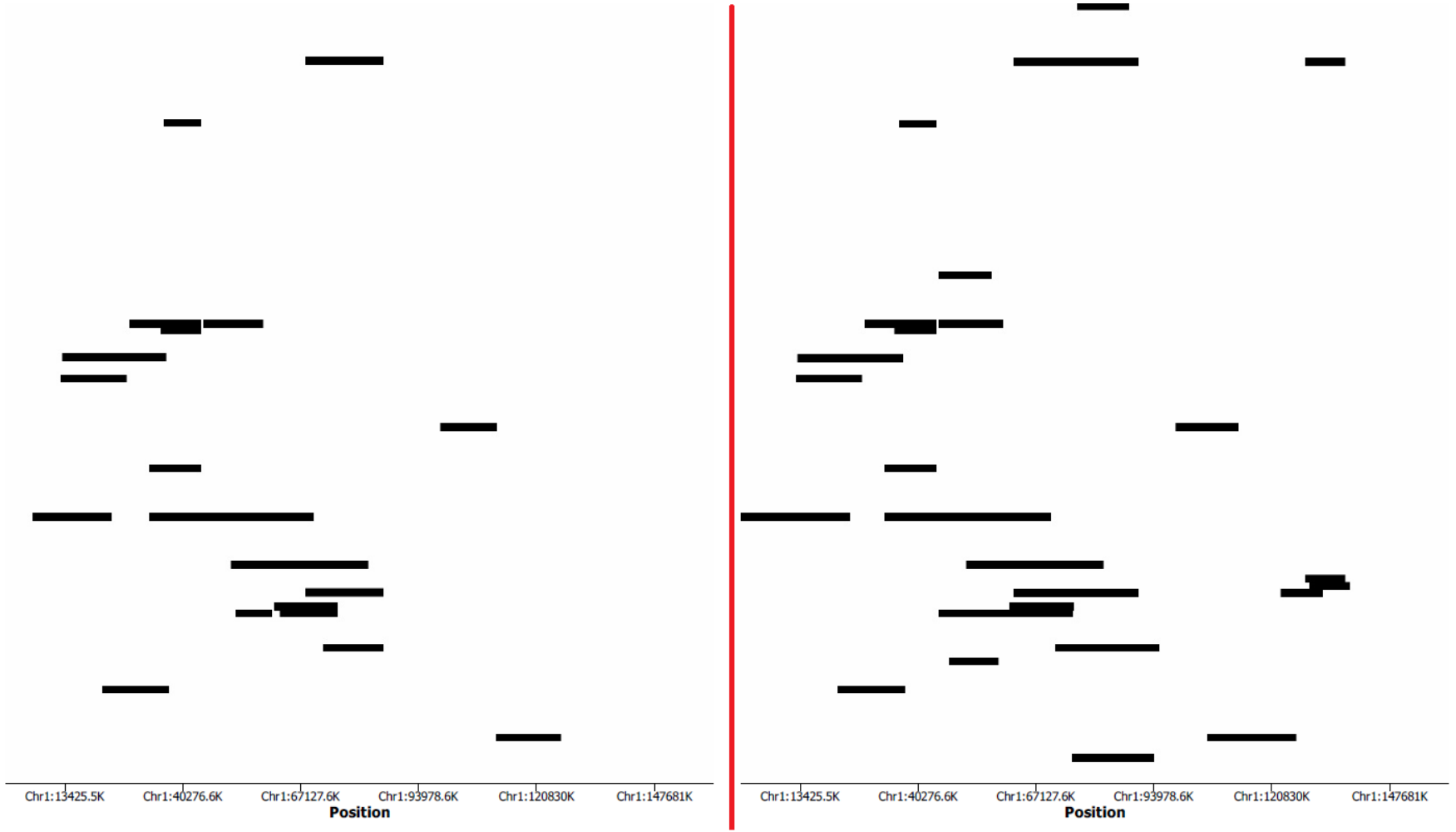# Importance of adjusting parameters



0    10

# Importance of adjusting parameters



Chr1:13425.5K  Chr1:40276.6K  Chr1:67127.6K  Chr1:93978.6K  Chr1:120830K  Chr1:147681K

**Position**

Chr1:13425.5K  Chr1:40276.6K  Chr1:67127.6K  Chr1:93978.6K  Chr1:120830K  Chr1:147681K

**Position**
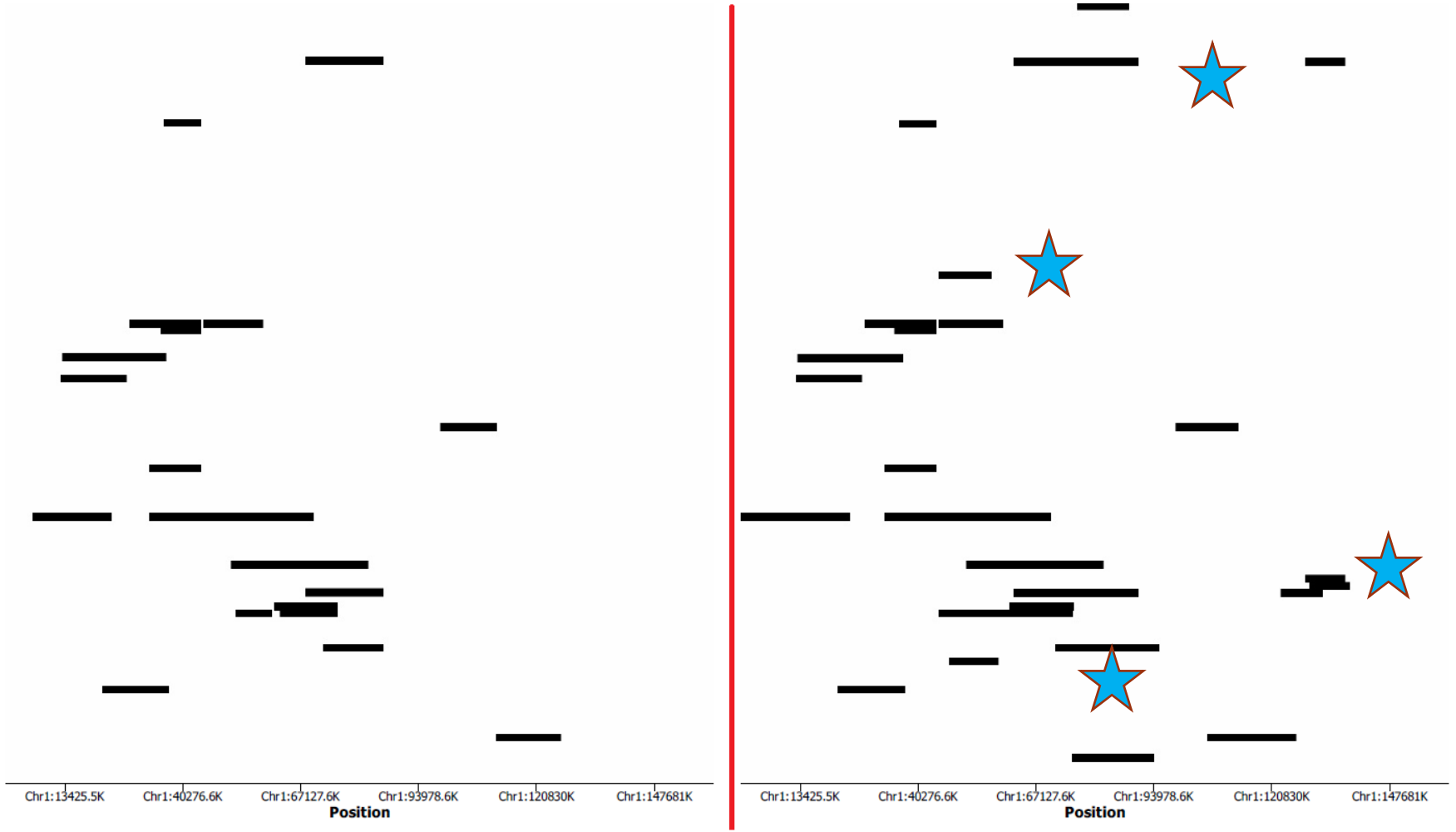
0                                                                    10

# Conclusions

- SNP panel density does affect estimation of ROH

- Adjustment of parameters is crucial for estimation

  - More work should be done on this field

- We recommend 54k panel for segments > 4 Mb

  - 54k panel overestimate $F_{ROH}$ for short segments – false IBD

- We recommend HD panel for all segments

  - HD panel can overestimate long segments – heterozygotes "in row"

    - Software developers should include restrictions in their algorithms

# Acknowledgments