Analysis of *de novo* & transmitted deletions in cleft case-parent trios

FaceBase Annual Meeting

May 1-3, 2013

Iowa City, IA

Younkin et al. (2013) A genome-wide study of *de novo* deletions identifies a candidate locus for non-syndromic isolated cleft lip/palate risk (submitted)

- PennCNV (Wang et al. 2007 Genome Res 11:1665) is based on a hidden Markov model (HMM) that jointly models the copy number states in parent-child trios (father, mother, child) using maximum likelihood methods
  - Widely used, but computationally intensive
- Minimum distance (Scharpf et al. 2012, BMC Bioinformatics 13:330) is specific for de novo deletions because it focuses on the difference in probe intensity (logR ratio & BAF) between child & each parent
  - Less likely to call a false positive by design
  - Computationally faster

Comparing CL/P & control trios (all of European ancestry)

- 479 CL/P case-parent trios from International Cleft Consortium
- 391 control trios from Dr. Marazita's study of dental caries
- Both studies genotyped on Illumina 660Quad chip at the Center for Inherited Disease Research (CIDR) as part of the GENEVA consortium

Detected *de novo* deletions in children

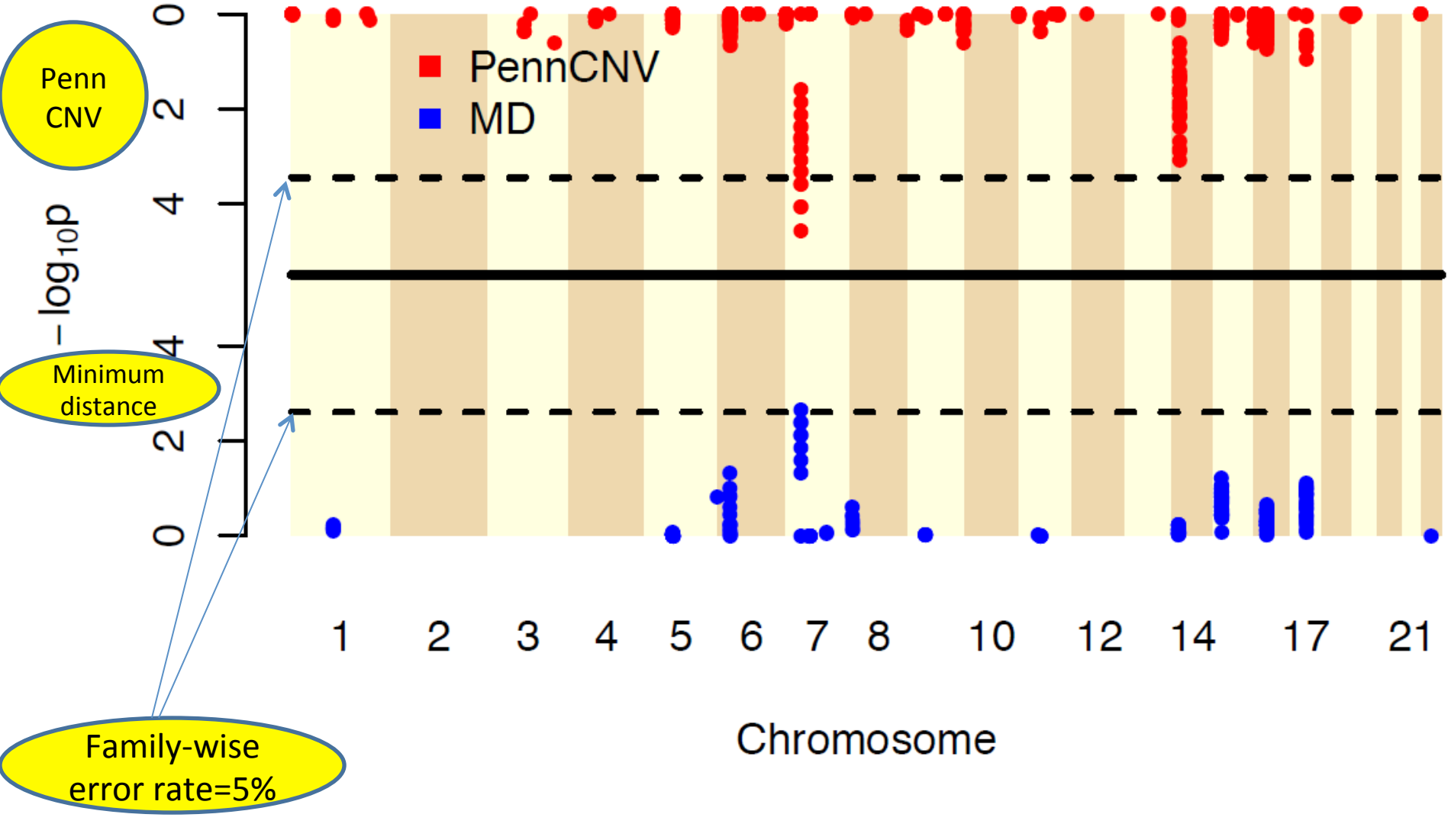
Counts	All controls		All cases		CLP cases		CL cases		CP cases	
	Count	Avg/ person	Count	Avg/ person	Count	Avg/ person	Count	Avg/ person	Count	Avg/ person
MinDist	454	0.85	287	0.43	117	0.46	90	0.56	80	0.51
PennCNV	1422	2.65	520	0.79	190	0.75	186	1.16	114	0.91

Width	All controls	All cases	CLP cases	CL cases	CP cases
MinDist	71.96	103.90	128.50	91.97	95.23
PennCNV	61.30	70.49	85.35	83.07	52.45

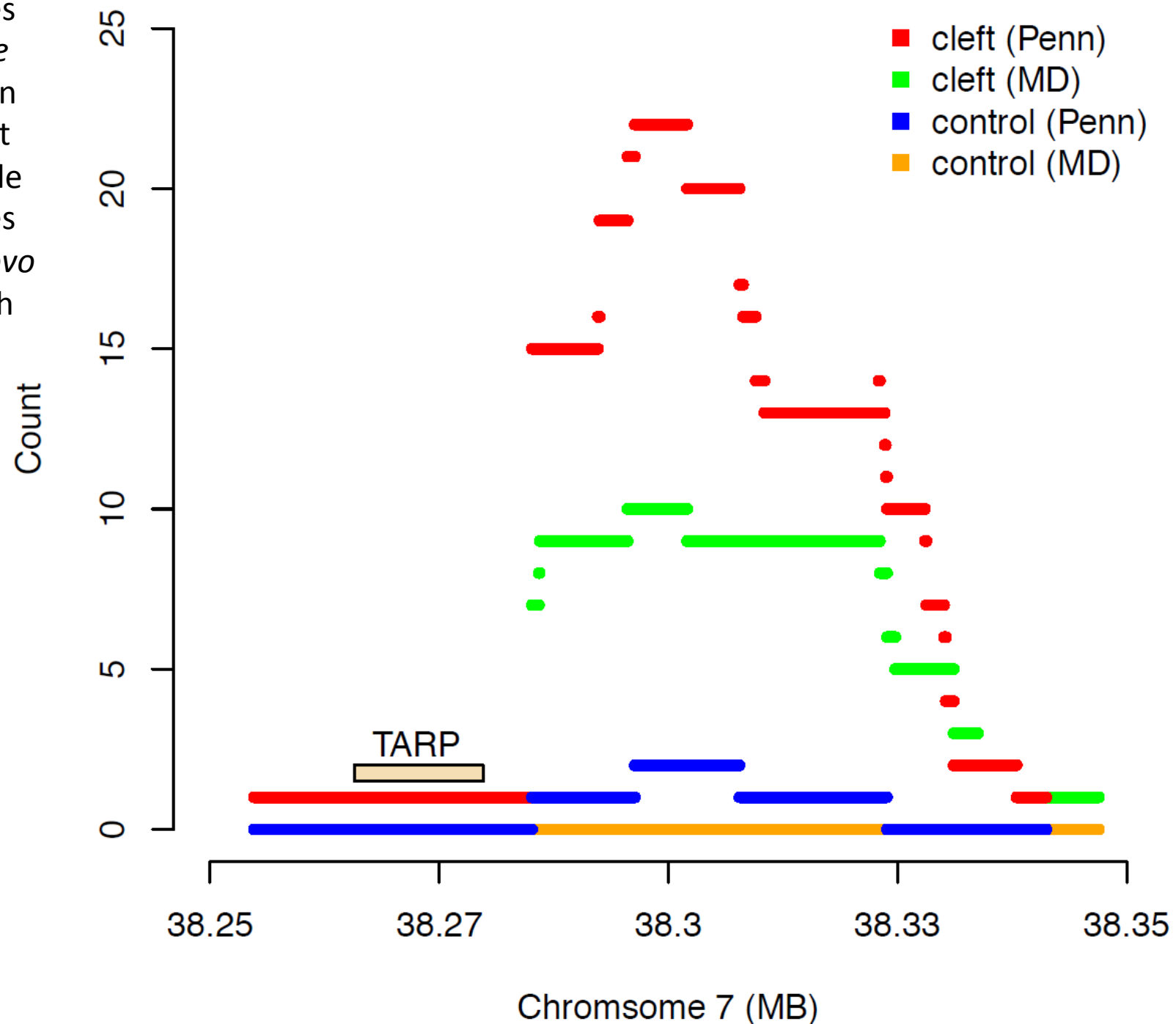
The difference in counts of CNVs may reflect differences in DNA source between cases and controls

- More control samples came from buccal swabs/saliva which creates noisier intensity data
- PennCNV tends to identify more short CNVs when data are noisy
- Minimum Distance is less likely to call a false positive *de novo* deletion

Figure 1: The -$\log_{10}$ p-values (y-axis) derived from testing associations of inferred de novo deletions and oral clefts, shown by chromosomal location (x-axis). Each point represents a de novo deletion CNV segment, delineated through *MinimumDistance* [43] (lower half) or *PennCNV* [42] (upper half). The dashed lines represent the genome-wide significance levels for a family-wise error rate of 5%, derived via permutation tests. The striped vertical bands indicate the 22 autosomes interrogated.



10 CL/P cases showed a *de novo* deletion with MinDist method, while 22 CL/P cases showed *de novo* deletion with PennCNV



Log(R ratio) & B allele frequency for 7p14.1

- Identified carriers have reduced log(R ratio)
- This region contains a number of structural variants

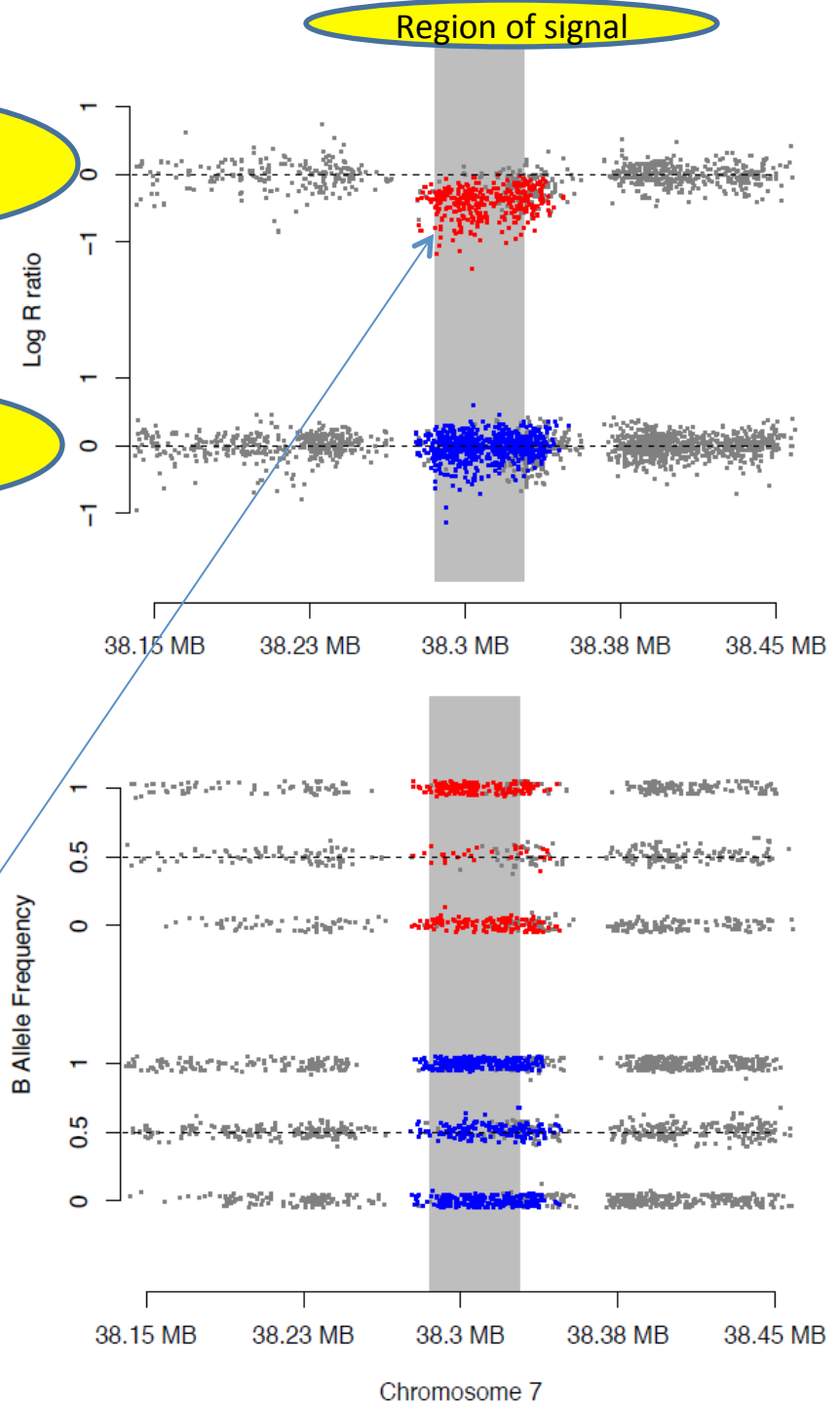
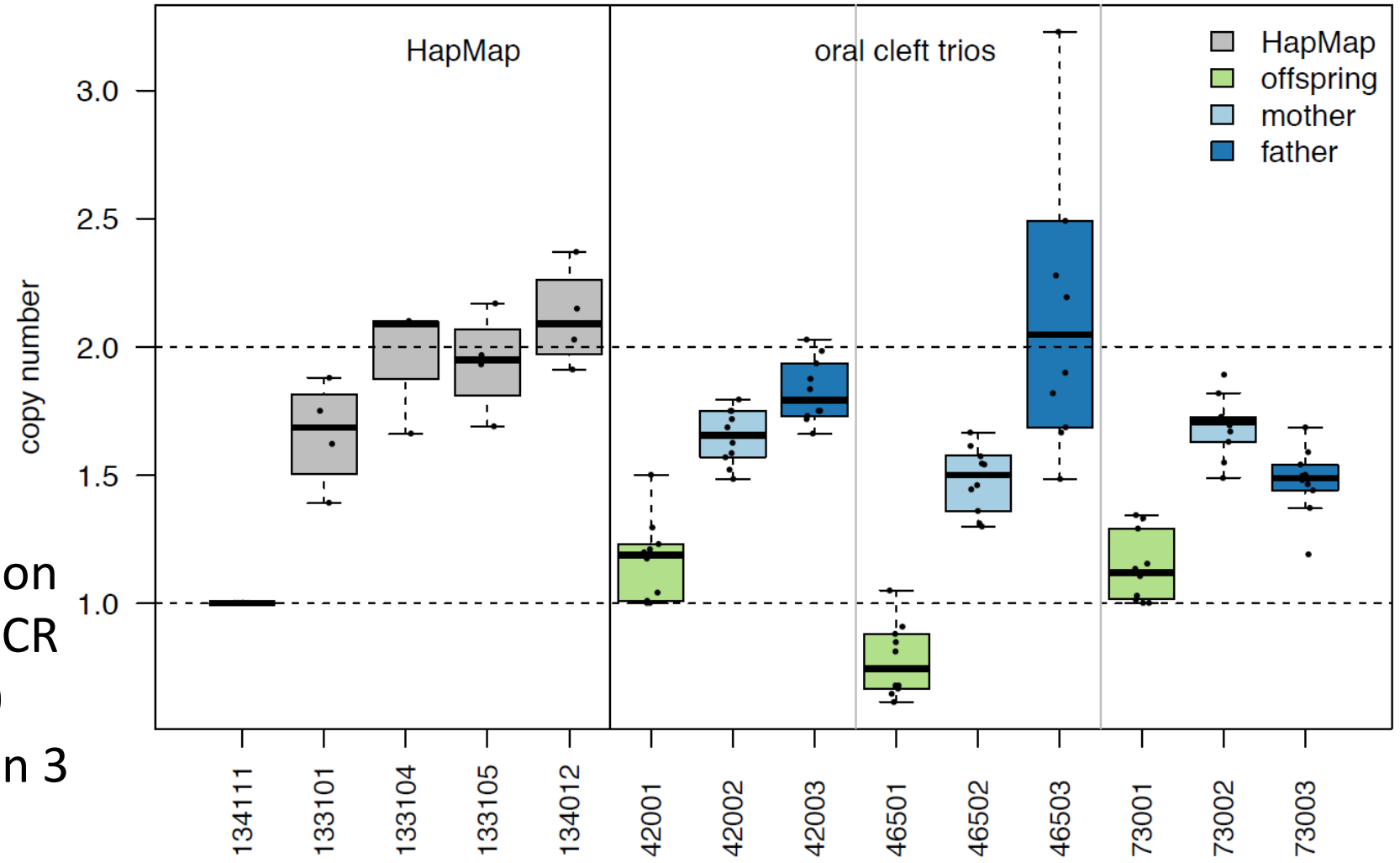


Figure S1: We used quantitative real-time PCR to validate 3 apparent de novo hemizygous deletions inferred from the Illumina arrays. Boxplots are used to summarize the copy number estimates (y-axis) obtained from 10 TaqMan probes located within the putative de novo deletion site for 14 samples (a boxplot for each sample). The samples include a positive control with an apparent hemizygous deletion (far left), 3 HapMap negative controls presumed to be diploid (boxplots 2-5), and 3 case-parent trios. The copy number estimate for each TaqMan probe was obtained using the CopyCaller software (v2.0) provided by the manufacturer with sample 134111 (HapMap id NA07034, far left) to calibrate hemizygous deletions.

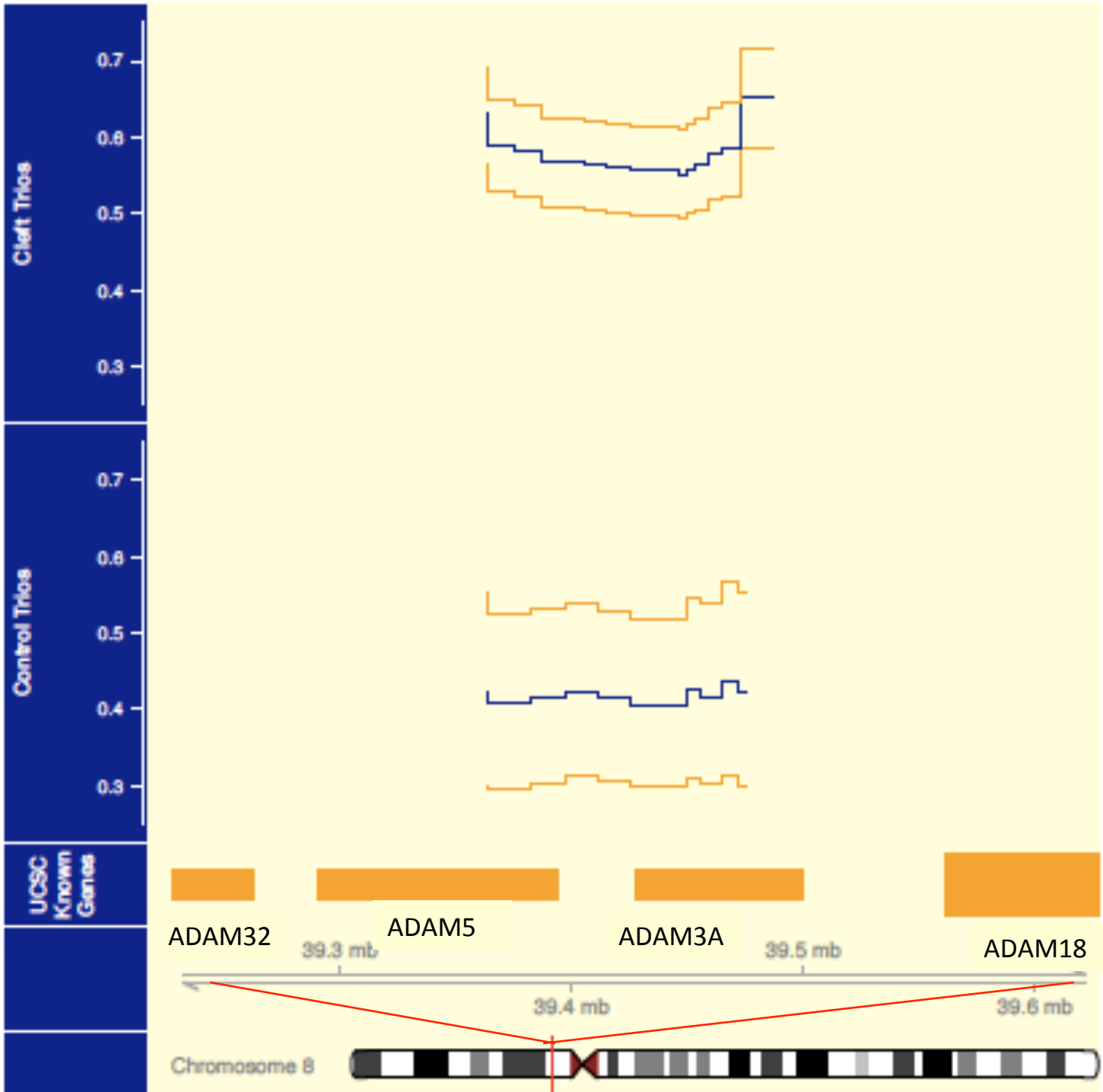


Validation with qPCR at 10 probes in 3 trios

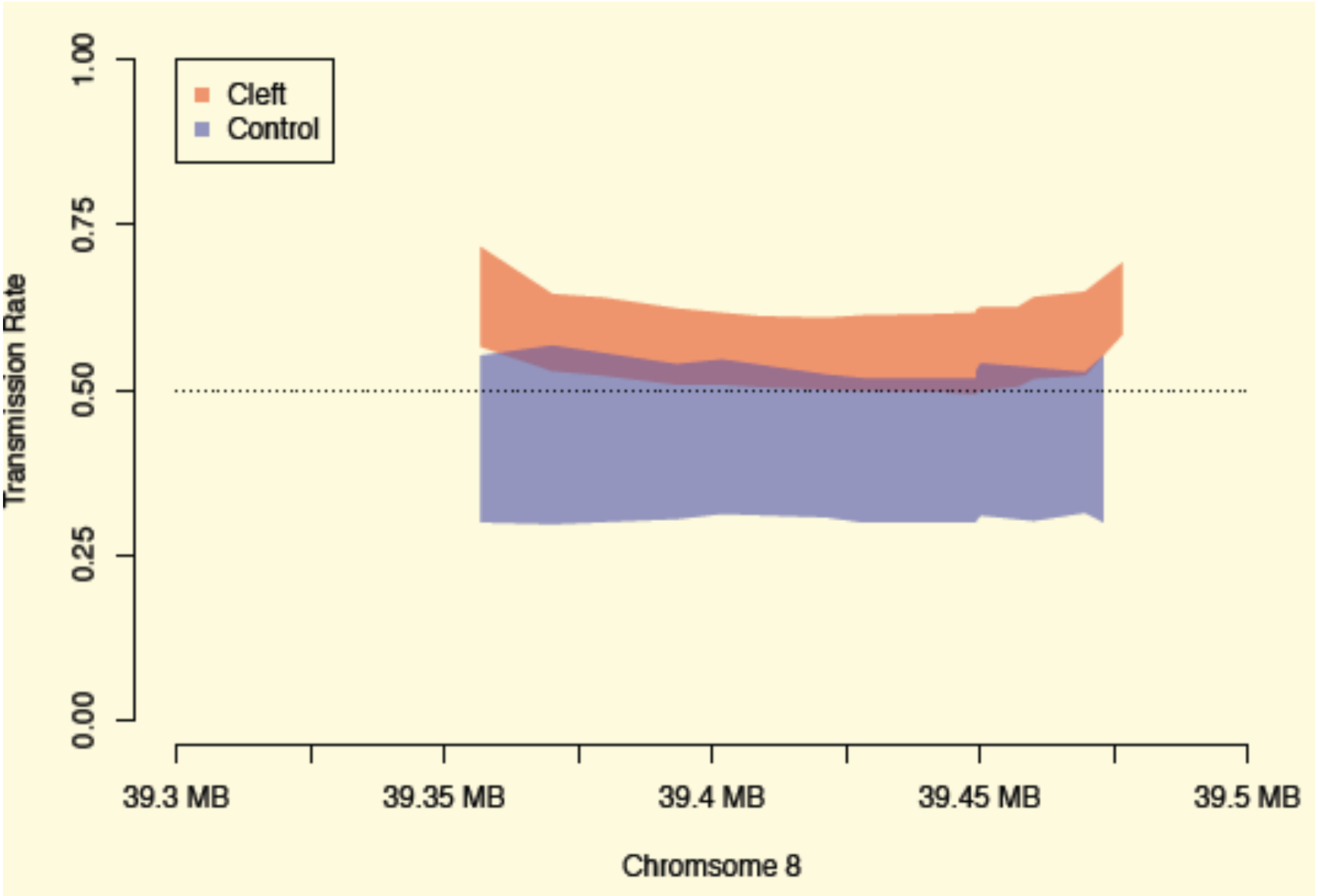
Preliminary results on transmitted CNVs in cleft and control trios

- Transmitted CNVs are based on PennCNV estimates
- All trios (Father, Mother, Child) were called for deletions over the genome
- A region harboring copy number polymorphisms (CNPs) on 8p11.23 showed possible over-transmission to children in these trios

Observed transmission rate (blue) & their 95%CI intervals (orange) for deletions in cleft case & control trios



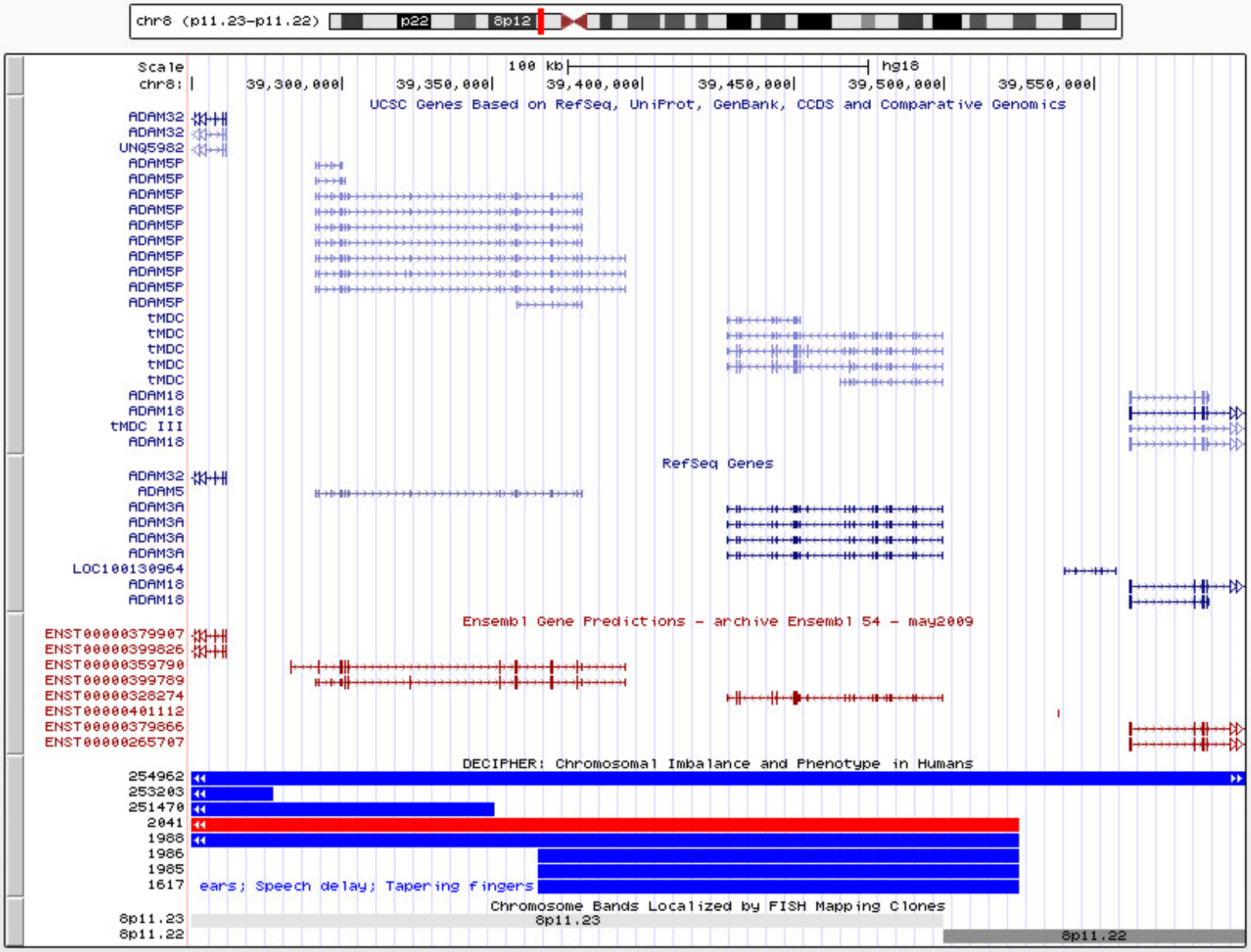
95%CI for estimated transmission of deletions in cleft trios on 8p11.23



UCSC Genome Browser on Human Mar. 2006 (NCBI36/hg18) Assembly

move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x

chr8:39,250,000-39,600,000 350,001 bp. enter position, gene symbol or search terms go



Possible issues with 8p11.23

- Sub-microscopic imbalances have been reported in this region
  - Associated with a range of phenotypic manifestations
- CNPs show different frequencies between parents of cleft cases and control parents
- The significance of the apparent over-transmission must be further explored
- Unequal recombination between these similar genes could also lead to apparent deletions in children

CNV calls from GWAS data on Facebase Hub

- PennCNV calls: submitted 6/7/12
- MinDist calls: submitted 2/8/13