

Figure 1: Cleft Trios — Red: $-\log_{10} p$ -values for each CNV component that has at least **five** “0/1” mating pairs. Null hypothesis is transmission rate of $\frac{1}{2}$, and alternative hypothesis is “one-sided.” Blue: The frequency of CNV component in the parents. Yellow: Gene tracks, labeled by transcript.

The top CNV component has the following trio-states. Where ‘1’ indicates a deletion and order is F, M, O.

```
000 001 010 011 100 101 110 111
149  7  26  57  28  71  5 102
```

The CNV components with significant p -values (Bonferroni) are given below.

GRanges with 16 ranges and 0 metadata columns:

	seqnames	ranges	strand
	<Rle>	<IRanges>	<Rle>
comp1994	chr8	[39356825, 39370186]	*
comp1995	chr8	[39370187, 39379683]	*
comp1996	chr8	[39379684, 39393417]	*
comp1997	chr8	[39393418, 39401865]	*
comp1998	chr8	[39401866, 39408899]	*
...
comp2005	chr8	[39450168, 39457081]	*

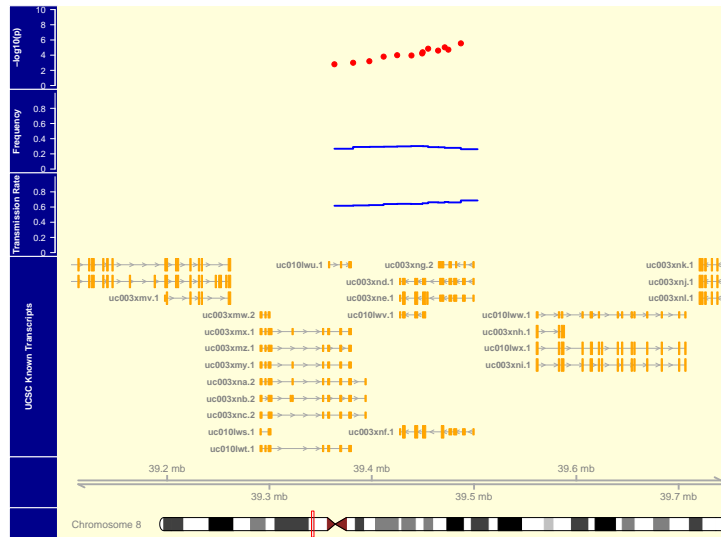


Figure 2: Control Trios — Red: $-\log_{10} p$ -values for each CNV component that has at least **five** “0/1” mating pairs. Null hypothesis is transmission rate of $\frac{1}{2}$, and alternative hypothesis is “one-sided.” Blue: The frequency of CNV component in the parents. Yellow: Gene tracks, labeled by transcript.

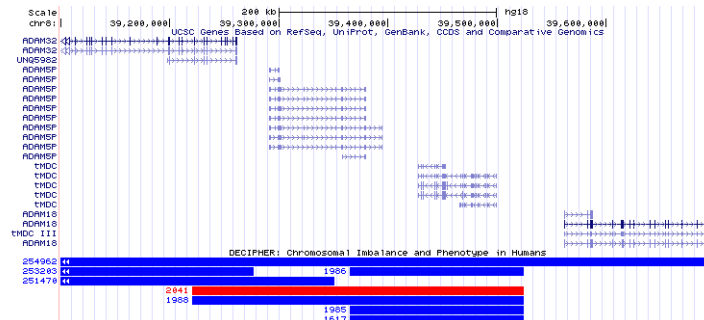


Figure 3: UCSC Genome Browser at chromosome 8 locus.

```

comp2006      chr8 [39457082, 39460087]      *
comp2007      chr8 [39460088, 39469612]      *
comp2008      chr8 [39469613, 39476658]      *
comp2009      chr8 [39476659, 39497557]      *
---
seqlengths:
      chr1      chr1_random      chr2 ...      chrY      chrM
      247249719      1663265      242951149 ...      57772954      16571

```

All CNV components are contiguous and the total width is 140.7 kB.

```

function (object)
{
  if (sum(object[c("101", "011", "100", "010")], na.rm = TRUE) >=
      5) {
    T <- sum(object[c("101", "011")], na.rm = TRUE)
    U <- sum(object[c("100", "010")], na.rm = TRUE)
    return(binom.test(x = T, n = T + U, p = 0.5, alternative = "greater")$p.value)
  }
  else {
    return(NA)
  }
}
<environment: namespace:trioClasses>

> gr.deletion.pitt <- gr.pitt[values(gr.pitt)$numsnp >= 10 &
  values(gr.pitt)$cn %in% 0:1]
> sum(countOverlaps(gr.deletion.pitt, reduce(locus)))

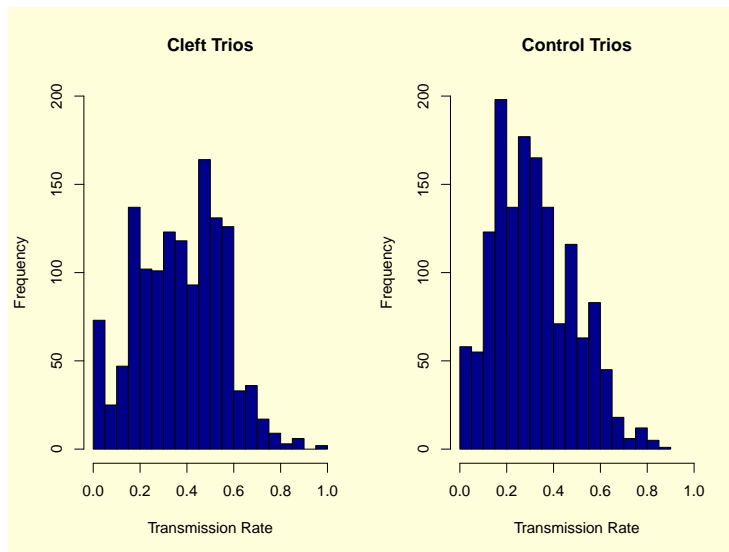
[1] 417

```

Methods

Cleft Data Description

- Performed 1346 tests. Bonferroni significant locus has width 140.733 kB.
- *PennCNV* joint HMM
- european, MAD < 0.3, non-WGA, aux \neq 1
- coverage > 10
- 12615 hemi/homozygous deletions identified in 445 trios



- 4288 CNV components
 - Common (> 0.01): 954
 - Rare: 3334
- Construct trio-states for all CNV components
 - recall that we use indicator variable for hemi/homozygous deletions
- must be at least 5 informative mating pairs
 - 01x and 10x
- count transmissions and perform `binom.test` (See “trans.tab”)

```
> mean(trans.rate[subjectHits(findOverlaps(reduce(locus), cnv.beaty.obj$cmp.gr))])
[1] 0.6669928

> mean(trans.rate.pitt[subjectHits(findOverlaps(reduce(locus),
  cnv.pitt.obj$cmp.gr))])
[1] 0.6474666

> mean(-log10(trans.vec[subjectHits(findOverlaps(reduce(locus),
  cnv.beaty.obj$cmp.gr))]))
[1] 5.870225

> mean(-log10(trans.vec.pitt[subjectHits(findOverlaps(reduce(locus),
  cnv.pitt.obj$cmp.gr))]))
```

```
[1] 4.164329
```

```
> freq.pitt.vec[subjectHits(findOverlaps(reduce(locus), cnv.pitt.obj$cmp.gr))]
```

```
comp1820 comp1821 comp1822 comp1823 comp1824 comp1825 comp1826  
0.2675781 0.2910156 0.2929688 0.2949219 0.2988281 0.3027344 0.3027344  
comp1827 comp1828 comp1829 comp1830 comp1831 comp1832  
0.3007812 0.2890625 0.2871094 0.2792969 0.2792969 0.2617188
```

```
> freq.vec[subjectHits(findOverlaps(reduce(locus), cnv.beaty.obj$cmp.gr))]
```

```
comp1994 comp1995 comp1996 comp1997 comp1998 comp1999 comp2000  
0.4445689 0.4960806 0.5039194 0.5139978 0.5139978 0.5162374 0.5207167  
comp2001 comp2002 comp2003 comp2004 comp2005 comp2006 comp2007  
0.5218365 0.5229563 0.5251960 0.5218365 0.5139978 0.5083987 0.4848824  
comp2008 comp2009  
0.4792833 0.4232923
```