

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
326  3  14  35  14  32   1  20
```

## Methods

### Cleft Data Description

- Performed 284 tests. Bonferroni significant locus has width 4.7 kB.
- *PennCNV* joint HMM
- european, MAD < 0.3, non-WGA, aux  $\neq$  1
- coverage > 10
- 13140 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”)

```
> est.list.beaty <- lapply(binom.list.beaty, FUN = trioClasses::get.est)
> ci.list.beaty <- lapply(binom.list.beaty, FUN = trioClasses::get.ci)
> testable.beaty <- !is.na(ci.list.beaty)
> ci.mat.beaty <- matrix(unlist(ci.list.beaty[testable.beaty]),
  nrow = sum(testable.beaty), ncol = 2, byrow = TRUE)
> est.vec.beaty <- as(est.list.beaty, "numeric")[testable.beaty]
> cnv.beaty.obj$cmp.gr[testable.beaty]
```

GRanges with 284 ranges and 0 metadata columns:

	seqnames	ranges	strand
	<Rle>	<IRanges>	<Rle>
comp120	chr1	[103969301, 103977611]	*
comp121	chr1	[103977612, 103983731]	*
comp122	chr1	[103983732, 103984810]	*
comp123	chr1	[103984811, 103988830]	*
comp124	chr1	[103988831, 103989217]	*

```

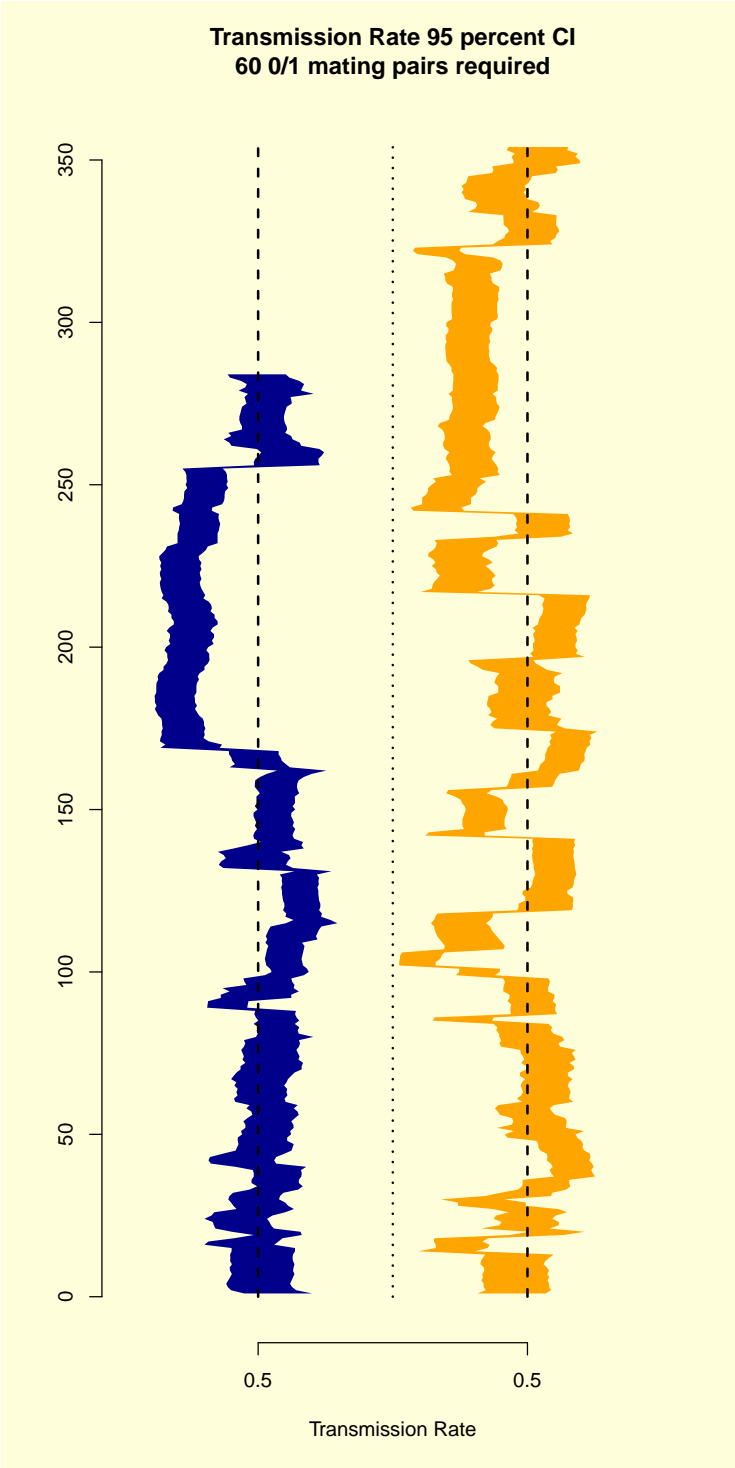
...      ...      ...      ...
comp3942  chr19    [20418019, 20422199]    *
comp3943  chr19    [20422200, 20493452]    *
comp3944  chr19    [20493453, 20507068]    *
comp4116  chr20    [52081230, 52081774]    *
comp4117  chr20    [52081775, 52088118]    *
---
seqlengths:
      chr1  chr1_random      chr2 ...      chrY      chrM
      247249719      1663265      242951149 ...      57772954      16571

> est.list.pitt <- lapply(binom.list.pitt, FUN = trioClasses:::get.est)
> ci.list.pitt <- lapply(binom.list.pitt, FUN = trioClasses:::get.ci)
> testable.pitt <- !is.na(ci.list.pitt)
> ci.mat.pitt <- matrix(unlist(ci.list.pitt[testable.pitt]),
      nrow = sum(testable.pitt), ncol = 2, byrow = TRUE)
> est.vec.pitt <- as(est.list.pitt, "numeric")[testable.pitt]
> cnv.pitt.obj$cmp.gr[testable.pitt]

GRanges with 354 ranges and 0 metadata columns:
      seqnames      ranges strand
      <Rle>      <IRanges> <Rle>
comp117  chr1 [103977612, 103983731]    *
comp118  chr1 [103983732, 103989830]    *
comp119  chr1 [103989831, 103990016]    *
comp120  chr1 [103990017, 103991756]    *
comp121  chr1 [103991757, 103999836]    *
...      ...      ...      ...
comp3774  chr20 [52081775, 52088118]    *
comp3831  chr21 [14163409, 14164422]    *
comp3832  chr21 [14164423, 14166029]    *
comp3833  chr21 [14166030, 14166821]    *
comp3834  chr21 [14166822, 14169701]    *
---
seqlengths:
      chr1  chr1_random      chr2 ...      chrY      chrM
      247249719      1663265      242951149 ...      57772954      16571

```

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



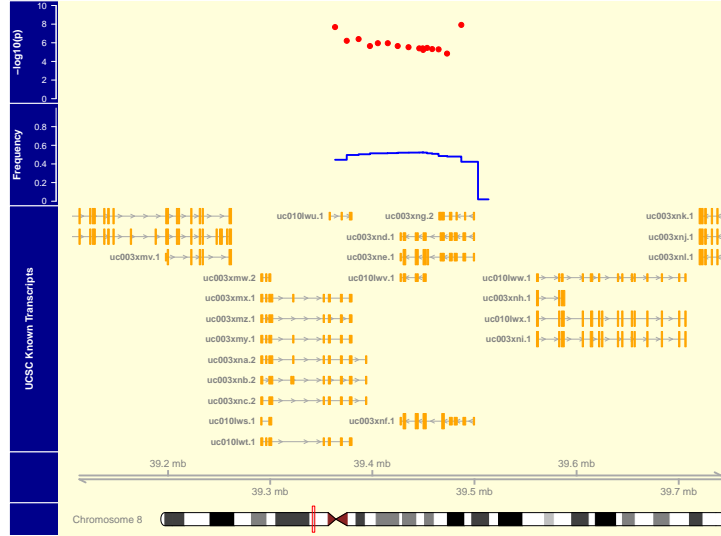


Figure 1: Cleft Trios — Red:  $-\log_{10} p$ -values for each CNV component that has at least **five** “0/1” mating pairs. Null hypothesis is transmission rate of  $\frac{1}{2}$ , and alternative hypothesis is “one-sided.” Blue: The frequency of CNV component in the parents. Yellow: Gene tracks, labeled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

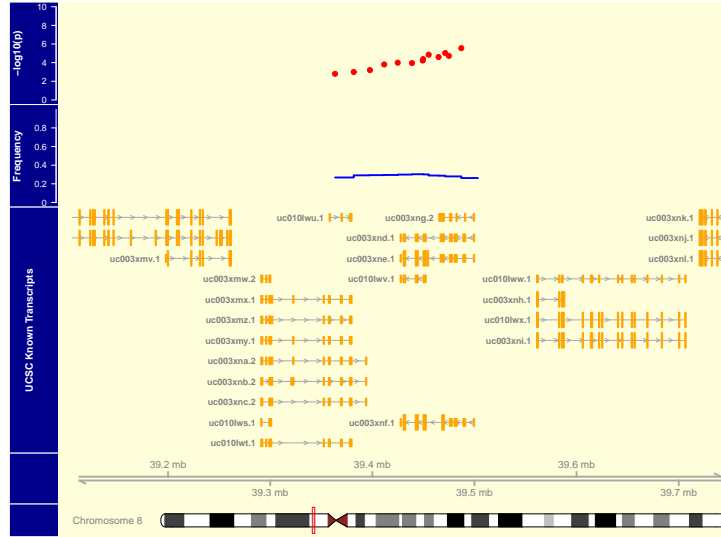
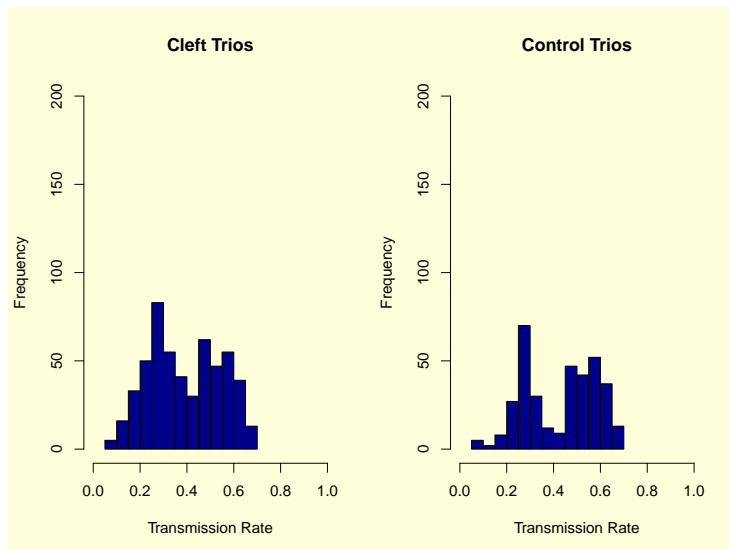


Figure 2: Control Trios — Red:  $-\log_{10} p$ -values for each CNV component that has at least **five** “0/1” mating pairs. Null hypothesis is transmission rate of  $\frac{1}{2}$ , and alternative hypothesis is “one-sided.” Blue: The frequency of CNV component in the parents. Yellow: Gene tracks, labeled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





```
<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] 649

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

[1] NA

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

[1] NA

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

[1] 5.784253

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

[1] 3.907709

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

```

comp1670 comp1820 comp1821 comp1822 comp1823 comp1824 comp1825
0.1914062 0.2675781 0.2910156 0.2929688 0.2949219 0.2988281 0.3027344
comp1826 comp1827 comp1828 comp1829 comp1830 comp1831 comp1832
0.3027344 0.3007812 0.2890625 0.2871094 0.2792969 0.2792969 0.2617188

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

comp1883 comp1994 comp1995 comp1996 comp1997 comp1998 comp1999
0.1534155 0.4445689 0.4960806 0.5039194 0.5139978 0.5139978 0.5162374
comp2000 comp2001 comp2002 comp2003 comp2004 comp2005 comp2006
0.5207167 0.5218365 0.5229563 0.5251960 0.5218365 0.5139978 0.5083987
comp2007 comp2008 comp2009
0.4848824 0.4792833 0.4232923

```