

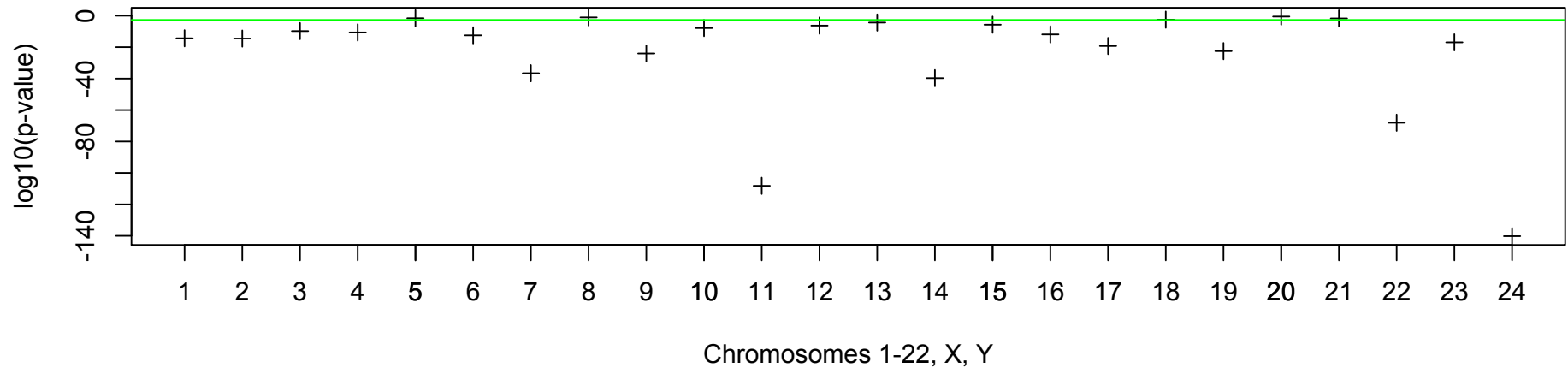
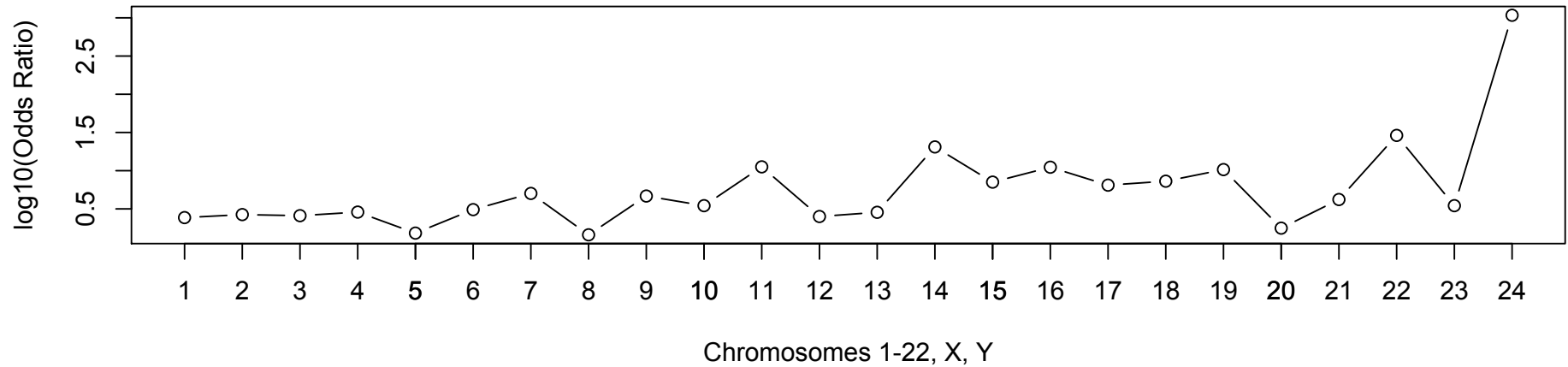
Human Pseudogenes' Localization Relative to their Parent Genes

Yan Zhang

6/25/2013

Pseudogene-parent Co-residence on Each Chromosome (Hyp5)

Fisher's Exact Test: Co-residence is Random or Not for Each Chromosome



Pseudogene-parent Co-residence on Each Chromosome (Hyp5)

Chromosome	Co-reside	OnlyPgeneOn	OnlyParentOn	Non	X-squared statistic	X-squared p-value	Fisher estimate of odds ratio	Fisher p-value
1	128	792	508	7644	73.65140086	9.32E-18	2.431629756	4.41E-15
2	107	636	495	7834	77.40672169	1.39E-18	2.662174542	2.85E-15
3	72	560	401	8039	51.13938452	8.60E-13	2.577123297	1.70E-10
4	64	517	351	8140	57.42892283	3.50E-14	2.870321118	2.33E-11
5	36	530	363	8143	5.041592608	0.02474573	1.523618257	0.025462278
6	68	483	371	8150	69.97502326	6.01E-17	3.092152214	3.26E-13
7	123	473	416	8060	243.7322861	6.04E-55	5.036668425	2.58E-37
8	25	440	325	8282	2.629979966	0.104863037	1.447820706	0.08352185
9	83	407	360	8222	159.3486245	1.57E-36	4.656361454	9.04E-25
10	33	344	233	8462	44.72290171	2.27E-11	3.483121676	1.44E-08
11	216	414	375	8067	852.4562939	2.13E-187	11.21901319	6.21E-109
12	45	245	598	8184	31.01559966	2.56E-08	2.513312964	5.20E-07
13	22	242	272	8536	20.84739604	4.97E-06	2.852400274	4.47E-05
14	50	98	217	8707	490.0616994	1.38E-108	20.44568294	2.00E-40
15	11	64	213	8784	41.75449639	1.03E-10	7.083019679	1.99E-06
16	19	48	310	8695	111.1060376	5.61E-26	11.0972869	1.26E-12
17	48	126	495	8403	143.2100621	5.29E-33	6.464182637	5.31E-20
18	4	37	132	8899	13.8136306	0.000201866	7.285069489	0.003184039
19	41	71	474	8486	196.8029568	1.04E-44	10.33165732	2.90E-23
20	4	146	136	8786	0.626687922	0.428572815	1.769794985	0.296508681
21	4	110	77	8881	6.185199971	0.012882314	4.192427167	0.018760749
22	77	151	153	8691	910.5976022	4.88E-200	28.94354455	8.62E-69
X	81	552	341	8098	99.80691458	1.68E-23	3.483910371	1.15E-17
Y	87	138	5	8842	3220.170609	0	1080.744215	6.00E-141

Y chromosome as destination

Y chromosome as source

most significant

Pseudogene-parent Co-residence on Each Chromosome (Hyp5)

- Results of Hyp5:
 - Co-residence is not random, but have high tendency for most chromosomes, esp. Y.
 - The extent of tendency is not the same for each chromosome.

Exchange among X, Y and Autochromosomes (Hyp6)

		Parent (source)		
		Autochromosomes	X	Y
Pseudogene (destination)	Autochromosomes	7729	306	4
	X	549	81	1
	Y	105	33	87

- 3by3: Pearson's Chi-squared test & Fisher's Exact Test: p-value < 2.2e-16.
- 2by2 (X, Y only): Fisher's Exact Test: p-value < 2.2e-16.
Sample estimat: odds ratio \approx 208.
- Y chromosome very seldom exports to other chromosomes.

Contributors of Co-residence

- Pseudogene types: PSSD or DUP
- Pseudogene transcription: Yes or No
- Chromosome

- Logistic regression models considering all above
 - Inter-effect between pseudogene type and transcribed or not is not significant

Logistic Regression Model (Hyp3)

logit(Coresidence) ~ Type + Transcribed + ParentChrom

Coefficients:

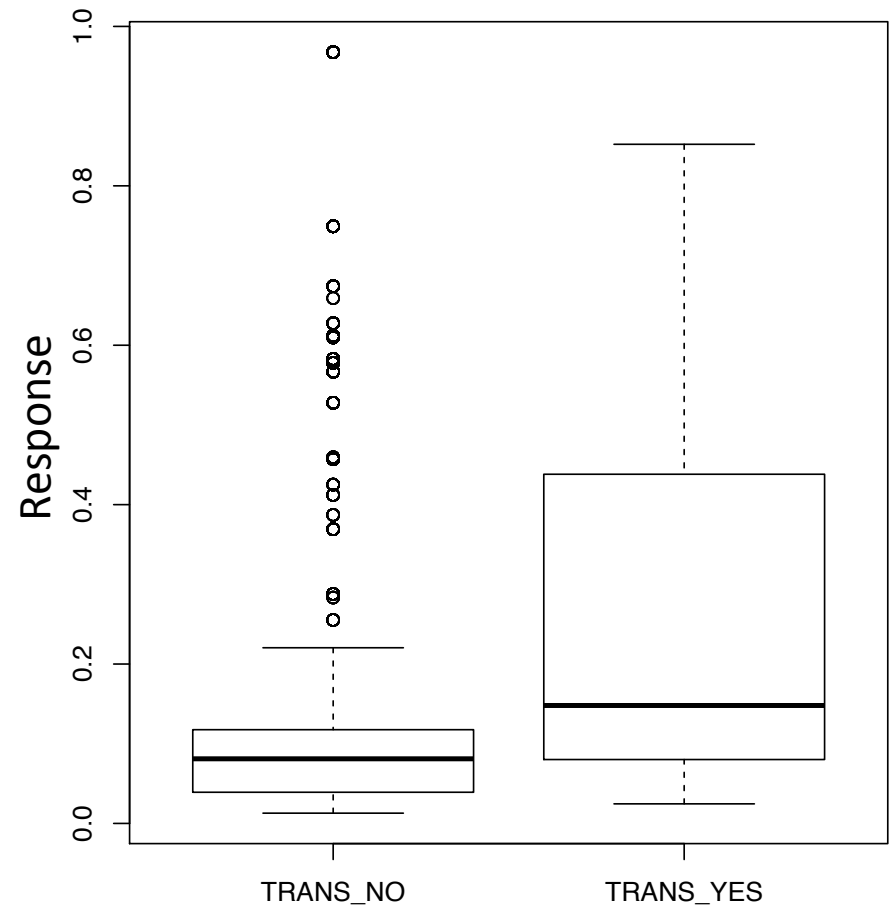
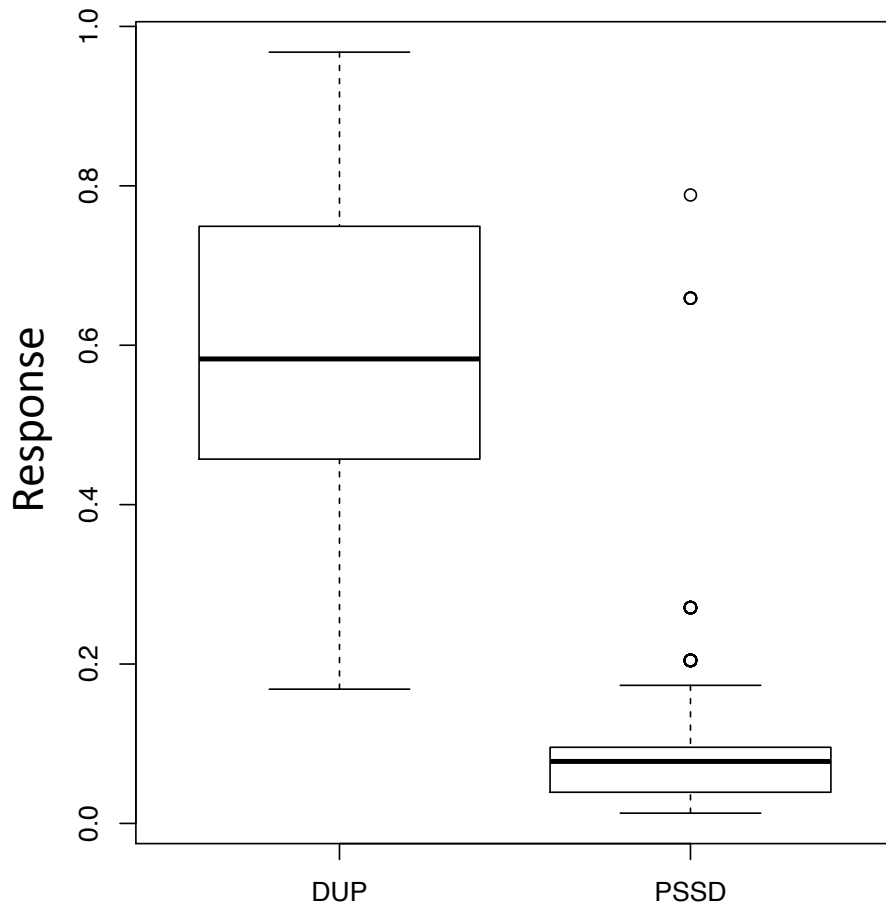
	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	7.266e-01	1.308e-01	5.554	2.79e-08	***
TypePSSD	-2.742e+00	7.867e-02	-34.850	< 2e-16	***
TranscribedTRANS_YES	6.568e-01	1.190e-01	5.519	3.40e-08	***
ParentChrom10	-8.986e-01	2.420e-01	-3.714	0.000204	***
ParentChrom11	3.680e-01	1.570e-01	2.344	0.019074	*
ParentChrom12	-1.028e+00	2.028e-01	-5.071	3.96e-07	***
ParentChrom13	-1.082e+00	2.662e-01	-4.064	4.83e-05	***
ParentChrom14	-2.078e+00	3.995e-01	-5.201	1.98e-07	***
ParentChrom15	-1.990e+00	3.543e-01	-5.616	1.95e-08	***
ParentChrom16	-1.797e+00	2.955e-01	-6.083	1.18e-09	***
ParentChrom17	-1.262e+00	2.049e-01	-6.161	7.24e-10	***
ParentChrom18	-2.323e+00	5.527e-01	-4.203	2.63e-05	***
ParentChrom19	-1.632e+00	2.237e-01	-7.298	2.93e-13	***
ParentChrom2	-4.125e-01	1.757e-01	-2.349	0.018844	*
ParentChrom20	-2.094e+00	5.422e-01	-3.862	0.000113	***
ParentChrom21	-1.654e+00	5.562e-01	-2.974	0.002939	**
ParentChrom22	-6.152e-01	2.338e-01	-2.632	0.008494	**
ParentChrom3	-2.043e-01	1.843e-01	-1.109	0.267627	
ParentChrom4	-4.578e-01	1.937e-01	-2.363	0.018124	*
ParentChrom5	-8.894e-01	2.269e-01	-3.920	8.86e-05	***
ParentChrom6	-2.714e-01	1.900e-01	-1.428	0.153151	
ParentChrom7	-2.798e-01	1.738e-01	-1.609	0.107516	
ParentChrom8	-1.186e+00	2.580e-01	-4.598	4.27e-06	***
ParentChrom9	-6.527e-05	1.803e-01	0.000	0.999711	
ParentChromX	-3.923e-01	1.856e-01	-2.113	0.034565	*
ParentChromY	2.674e+00	5.041e-01	5.305	1.13e-07	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Logistic Regression Model (Hyp3)

- Results from fitted model – whole genome

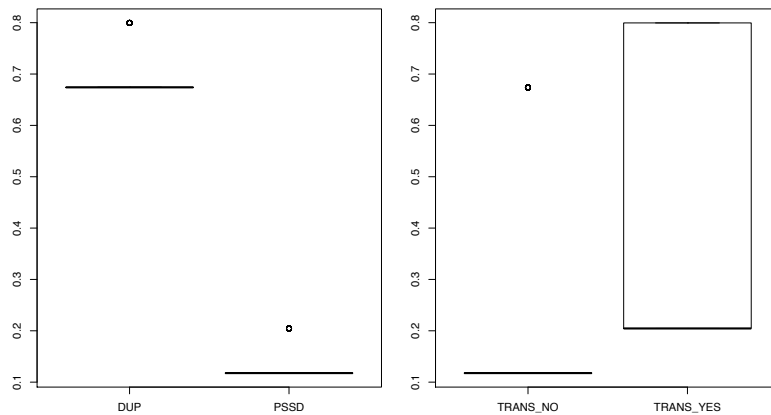
Whole genome -- calculated from fitted model



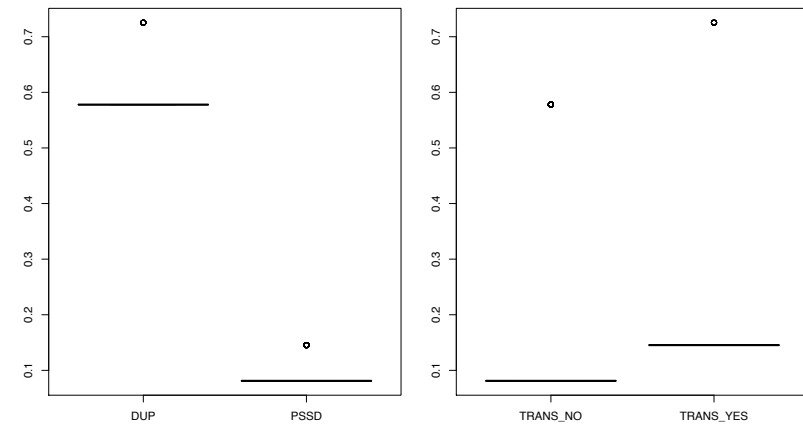
Logistic Regression Model (Hyp3)

- Results from fitted model – individual chromosomes, just examples shown:

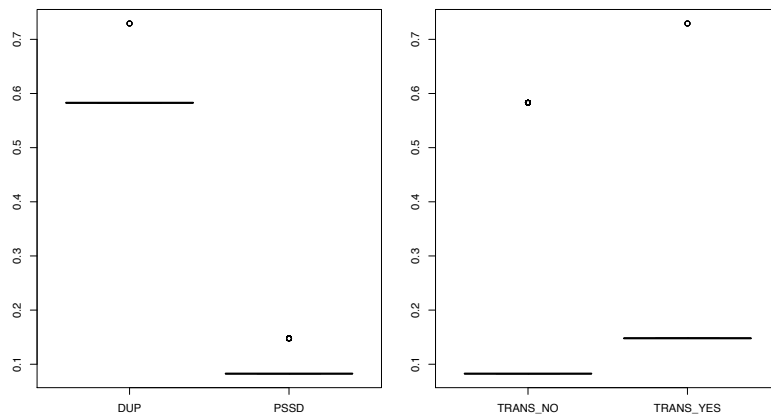
Chromosome 1 -- calculated from fitted model



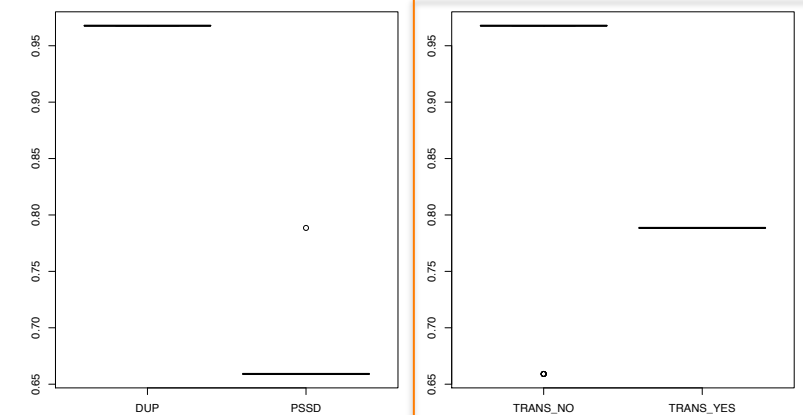
Chromosome 2 -- calculated from fitted model



Chromosome X -- calculated from fitted model



Chromosome Y -- calculated from fitted model



Only exception

Reference slide (1 slide from Hyp7)

- trPSSD-parent co-reside on Y, compared with background

Both on Chr Y	Only pseudogene on Chr Y	Only parent on Chr Y	Non on Chr Y
0	0	1	442

- trDUP-parent co-reside on Y, compared with background

Both on Chr Y	Only pseudogene on Chr Y	Only parent on Chr Y	Non on Chr Y
0	2	0	190

- nontrPSSD-parent co-reside on Y, compared with background

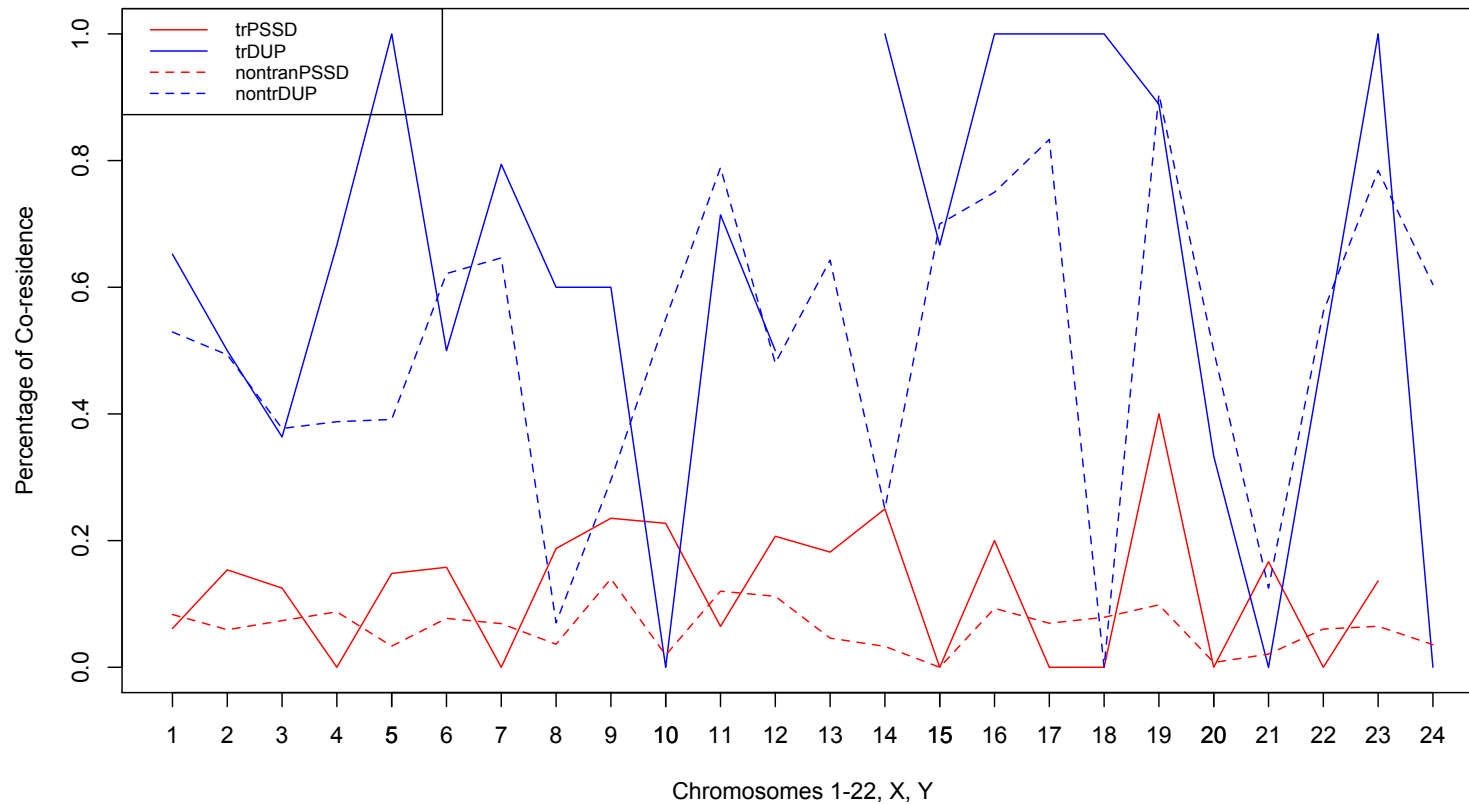
Both on Chr Y	Only pseudogene on Chr Y	Only parent on Chr Y	Non on Chr Y
3	81	3	7055

- nontrDUP-parent co-reside on Y, compared with background

Both on Chr Y	Only pseudogene on Chr Y	Only parent on Chr Y	Non on Chr Y
84	55	1	981

An Old Test (Hyp4): Four Subtypes vs. Co-residence/Not

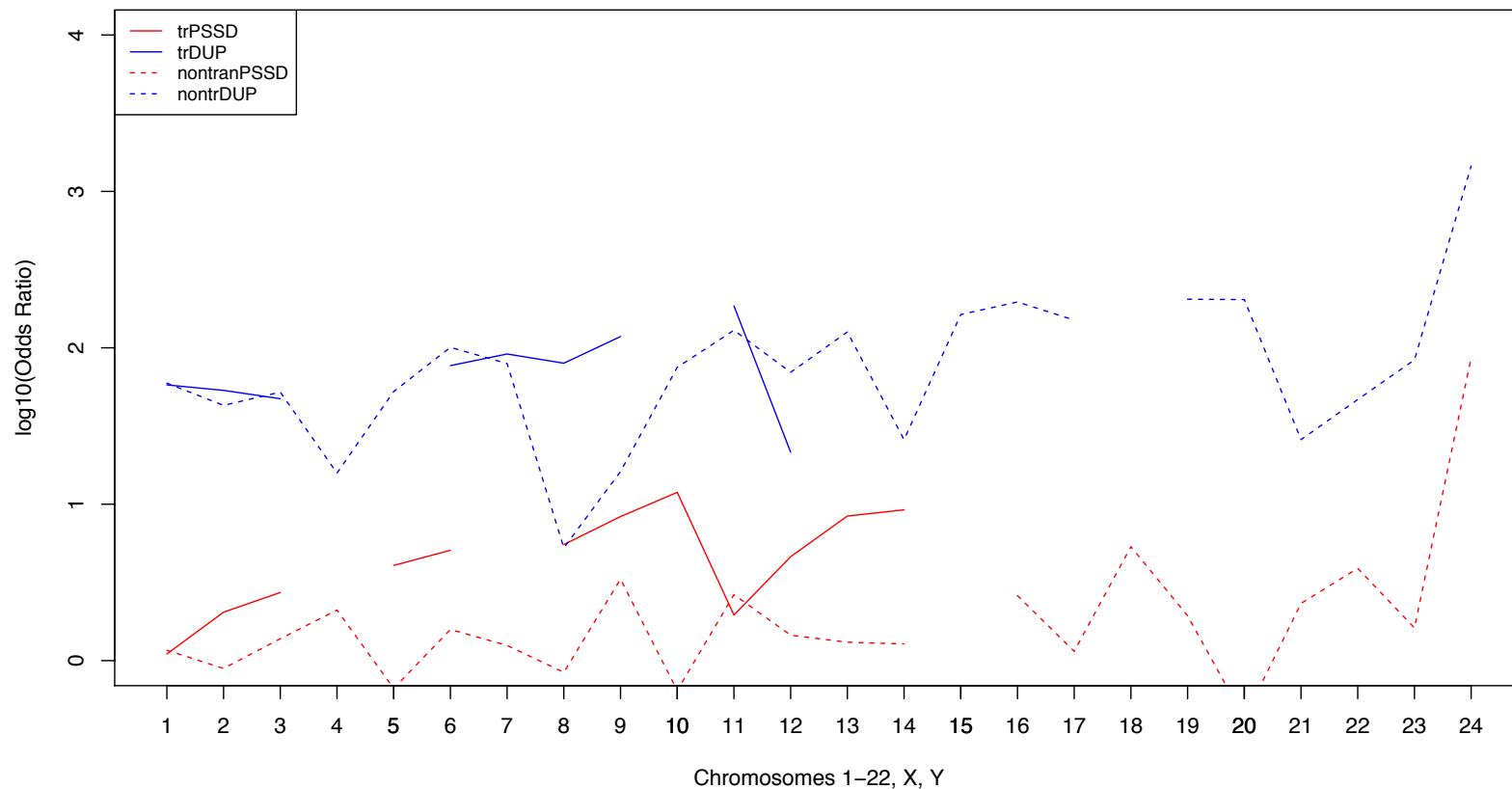
Percentage of Co-residence for 4 Subtypes



Note: Hyp4 Separate data by the Chrom the pseudogene is on. Not by the ParentChrom. Confounding effect of Chrom should still influence this result.

Hyp7: For four subtype, is co-residence random (compared with background) for each chromosome?

Fisher's Exact Test: Co-residence is Random or Not for Each Chromosome
Four Odds Ratio Lines



Log odds ratio is shown above. Significant or not depends on chromosome.