

# **Exchange Among Sex Chromosomes and Autosomes - Human (ctd)**

Yan Zhang

7/30/2013

# Outline

- Parent genes on Y
- Parent genes on X
- Linear regression

**Parent genes on Y**

# Some observations in human

As seen before:

- 14 Unique Human Parent Transcript IDs on Y

	ParentTranscriptID	Freq.ParentT	ParentGeneN	ParentStart	ParentEnd	ParentLength
1	ENST00000250784	1	RPS4Y1	2709527	2734997	25470
2	ENST00000288666	1	RPS4Y2	22918050	22942918	24868
3	ENST00000303728	4	PRY	24636544	24660784	24240
4	ENST00000304790	1	HSFY1	20933700	20935621	1921
5	ENST00000306609	11	CDY1	27768309	27771049	2740
6	ENST00000307393	3	HSFY1	20708557	20710478	1921
7	ENST00000331070	1	BPY2	25130410	25151606	21196
8	ENST00000336079	1	DDX3Y	15016742	15032390	15648
9	ENST00000338981	29	USP9Y	14813160	14972764	159604
10	ENST00000361963	9	CDY1	27768264	27770483	2219
11	ENST00000382707	21	RBMX1A1	23696765	23711212	14447
12	ENST00000421178	2	FAM197Y1	9374241	9384693	10452
13	ENST00000426950	1	TSPY4, 8, 10	9175073	9177887	2814
14	ENST00000451548	7	TSPY1	9304564	9307357	2793

Question: Are these high freq parent genes highly expressed, in specific tissues?

# Some observations in human

- High freq parent genes on Y:
  - CDY1, USP9Y, RBMY1A1 and TSPY1 have high number of pseudogenes.
  - All four (highly) expressed in testis, i.e. germline tissue (based on RNA-seq Body Map).
- Low freq parent genes on Y:
  - All are Y-linked.
  - All expressed in testis, except FAM197Y1, which is expressed in adrenal gland (RNA-seq data).

**Parent genes on X**

- From literature: “The Y chromosome has been shown to recruit male-specific genes, whereas a few individual X-linked genes have male specific duplicate counterparts on autosomes.” “Some male-specific genes appear to be enriched on the X chromosomes.” [1]
- Rough look of parent genes on X:
  - No Y-linked functions have been found enriched.

# High freq parent genes on X

ParentTranscriptID	Freq.Parent	ParentStart	ParentEnd	Gene		
ENST00000427805	36	100645812	100651105	RPL36A	Ribosomal Protein L36a	
ENST00000316084	22	71491892	71497150	RPS4X	Ribosomal Protein S4, X-Linked	
ENST00000361575	22	118920469	118925606	RPL39	Ribosomal Protein L39	
ENST00000424325	19	153626571	153630680	RPL10	Ribosomal Protein L10	
ENST00000325307	17	150151759	150159248	HMGB3	High Mobility Group Box 3	
ENST00000380567	15	13752864	13787472	OFD1	Oral-Facial-Digital Syndrome 1	
ENST00000317881	9	118602363	118605282			
ENST00000392994	9	100645977	100650781			
ENST00000481445	8	77154935	77162870			
ENST00000464506	7	13707244	13728625			
ENST00000406022	6	153627231	153629250			
ENST00000305536	5	129535943	129547317			
ENST00000404933	5	21958691	22012953			
ENST00000320676	4	135955620	135962884			
ENST00000328078	4	120181462	120183794			
ENST00000381401	4	1505045	1511617			
ENST00000217926	3	103265719	103268259			
ENST00000217958	3	107327437	107334848			
ENST00000218104	3	152990323	153010216			
ENST00000318433	3	71401526	71418187			

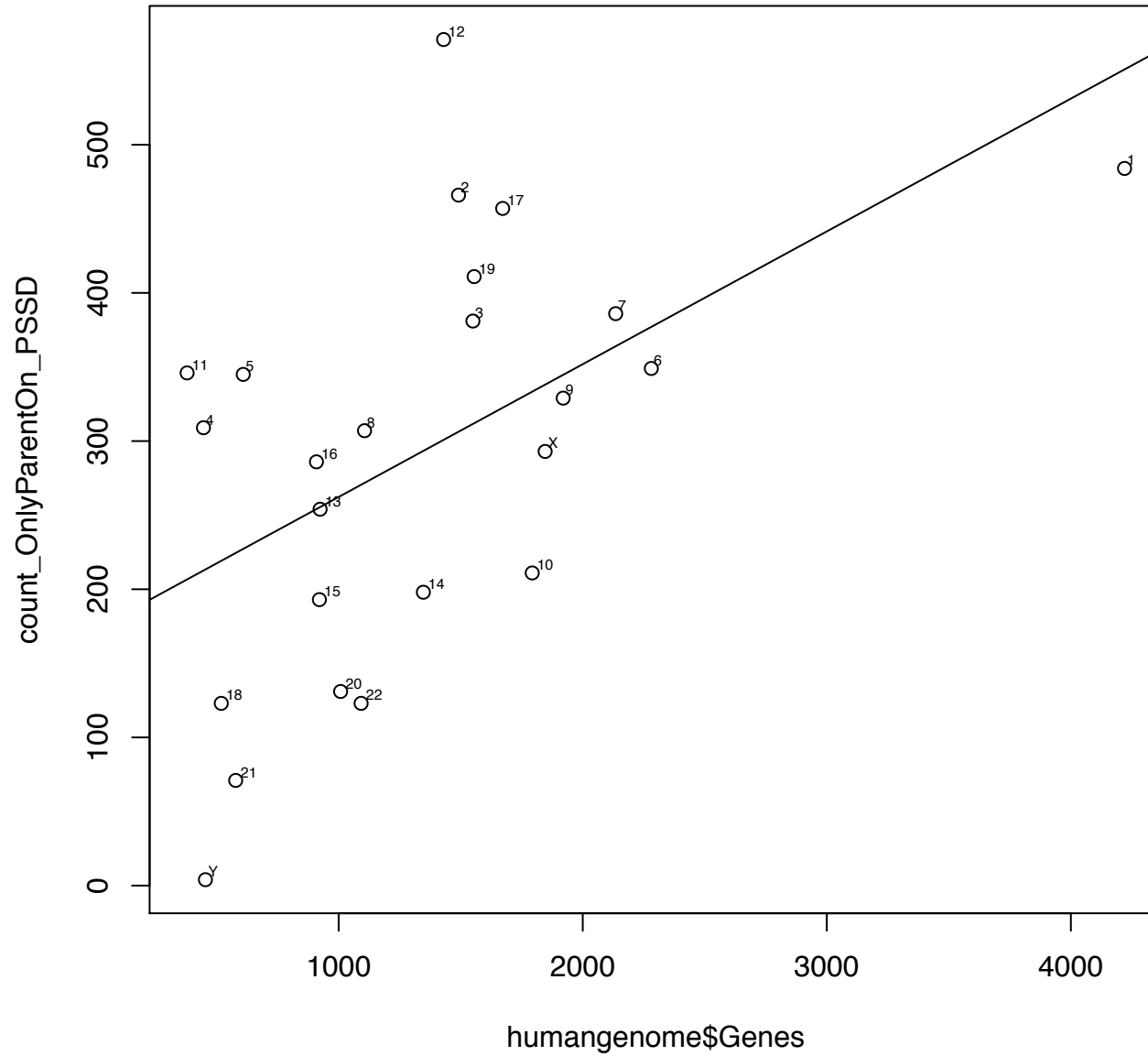


# Linear regression

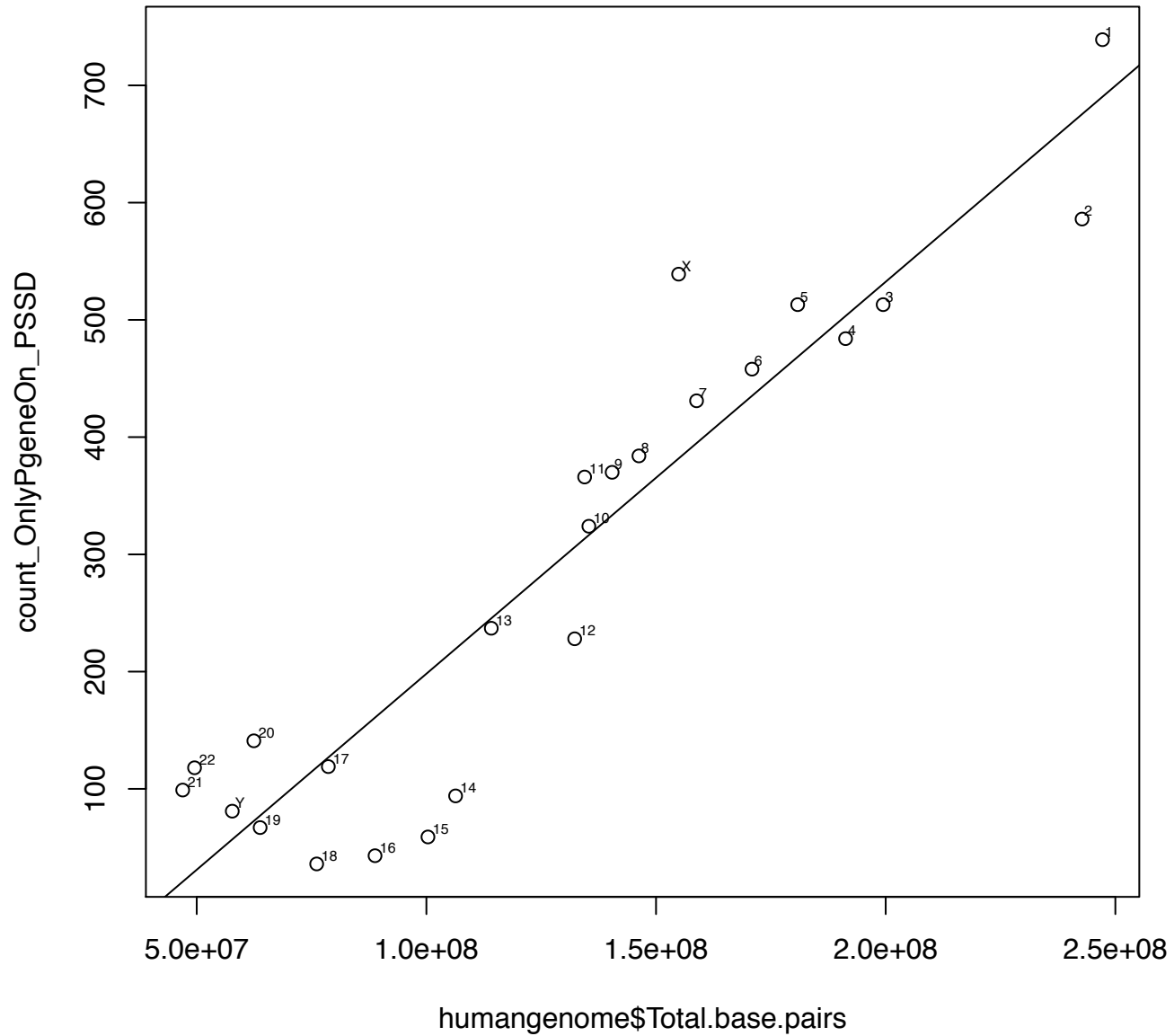
# Regressions

- #Genes vs. #Parents
- Size of chromosomes vs. #Pseudogenes, PSSD or DUP

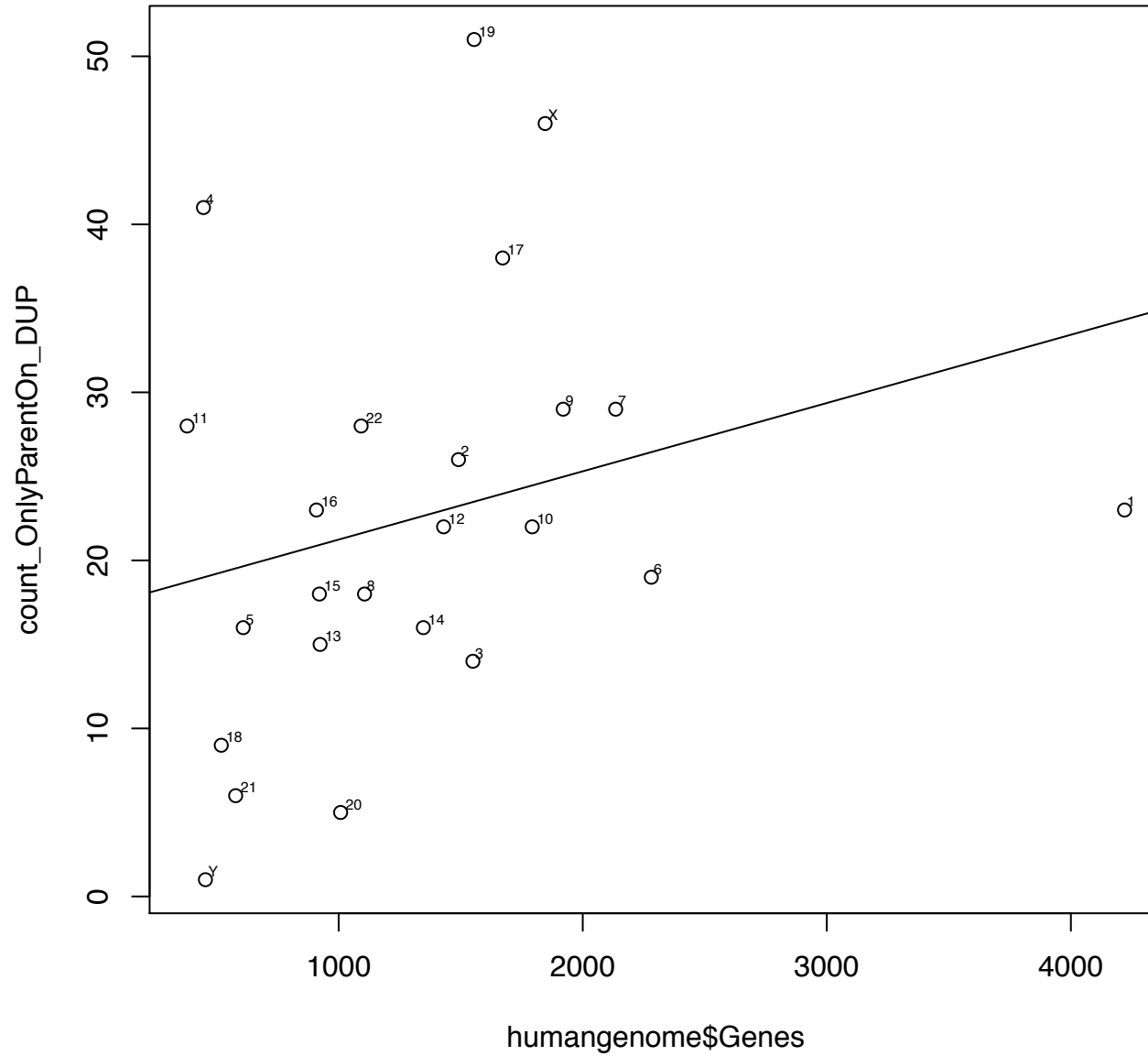
# PSSD



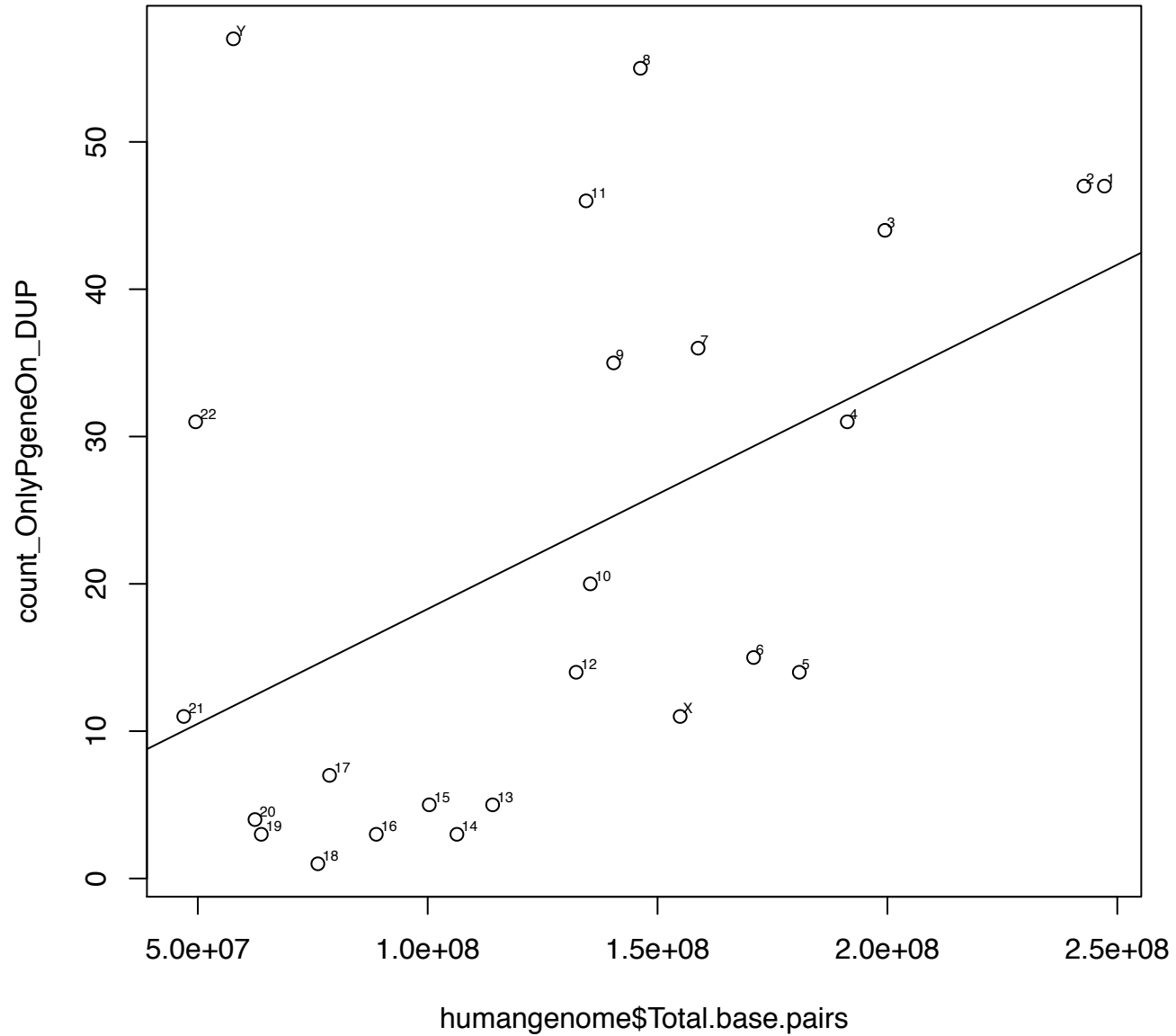
# PSSD



# DUP



# DUP



# Reference:

## Rough summary from their work (ctd)

- Linear regression results (Y was not included):

Fig. 1. Regressions for the parental genes of retrogenes in (A) human and (C) mouse and for the parental genes of retropseudogenes (B) in human. Regressions for the size of a chromosome in (D) human and (F) mouse and for the retropseudogenes (E) in human. In the plots, X is shown as 75% of its size as predicted by the model (15), although allowing X to assume 100% of its size does not change the results. Probabilities for the hypothesis that the chromosome with the highest observed/expected ratio [where the expected number is calculated as in (16)] is an outlier are calculated using Grubbs and Dixon outlier tests (16). For every distribution [except (B)], the X has the largest ratio and is an outlier with  $P < 0.005$  and  $P < 0.01$  for the Grubbs and Dixon tests, respectively; (B) shows no such outliers.

