

# GATK Unified Genotyper on chrX and chrY

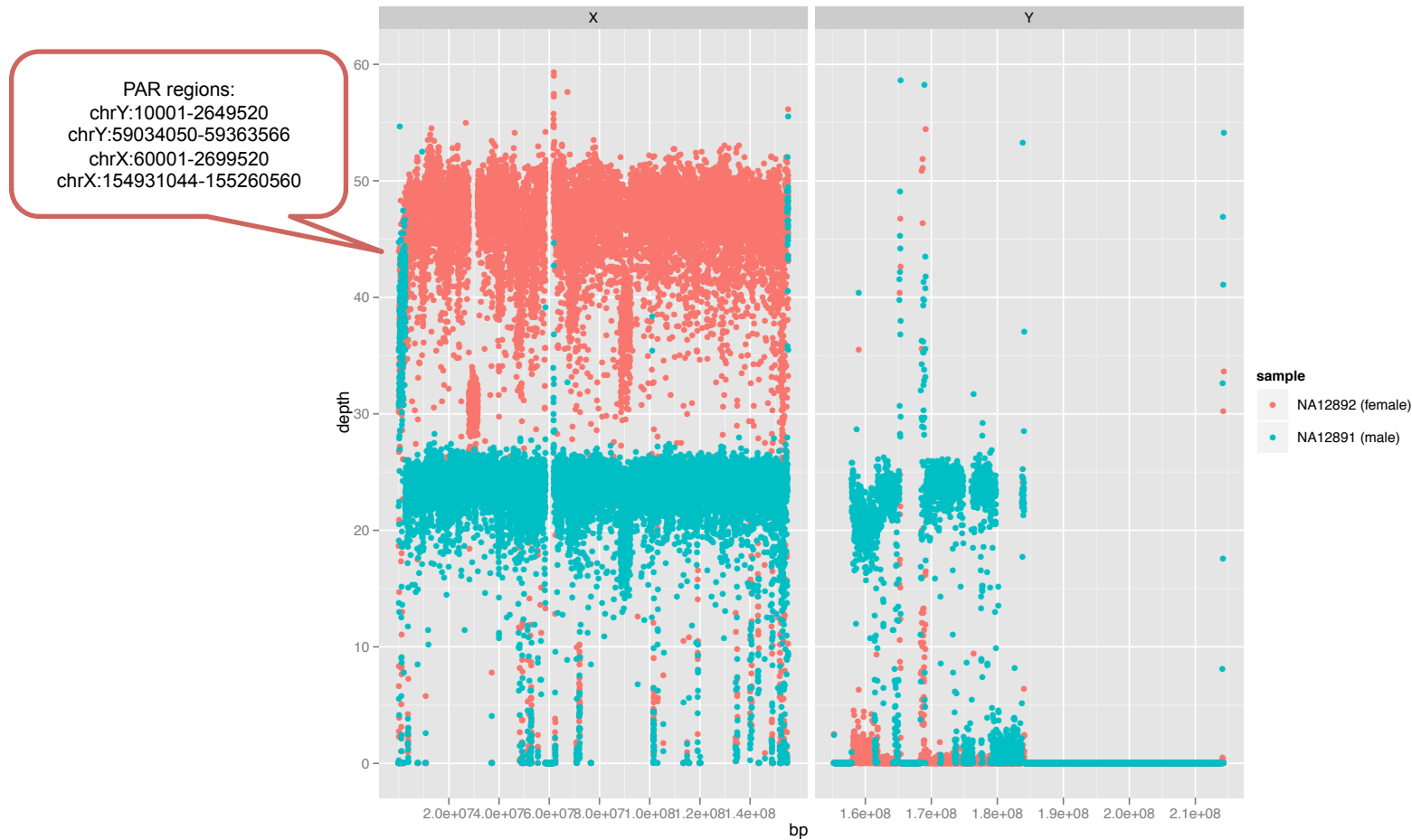
Mark DePristo

4/6/11

# Data and definitions

- NA12891 and 92 (CEU dad and mom) on HiSeq to ~30x coverage
- Indel calling at known sites only
  - Not fully indel cleaned as per best practices
- UG calls unfiltered
  - How good are the raw calls?

# Coverage looks good on sex chromosomes



# How many calls are made on each sex chromosome by gender?

Sample	Gender	Chrom	No. hets	No. hom-var
Unfiltered				
NA12891	Male	X	11273	75152
NA12892	Female	X	78276	41635
NA12891	Male	Y	14010	2090
NA12892	Female	Y	7876	7296
Filtered				
NA12891	Male	X	1609	69824
NA12892	Female	X	57504	38207
NA12891	Male	Y	1581	1423
NA12892	Female	Y	582	2573

Excluding PAR regions

Simple hard filters: QD < 5.0 || SB > -0.10 || HaplotypeScore > 3.0