

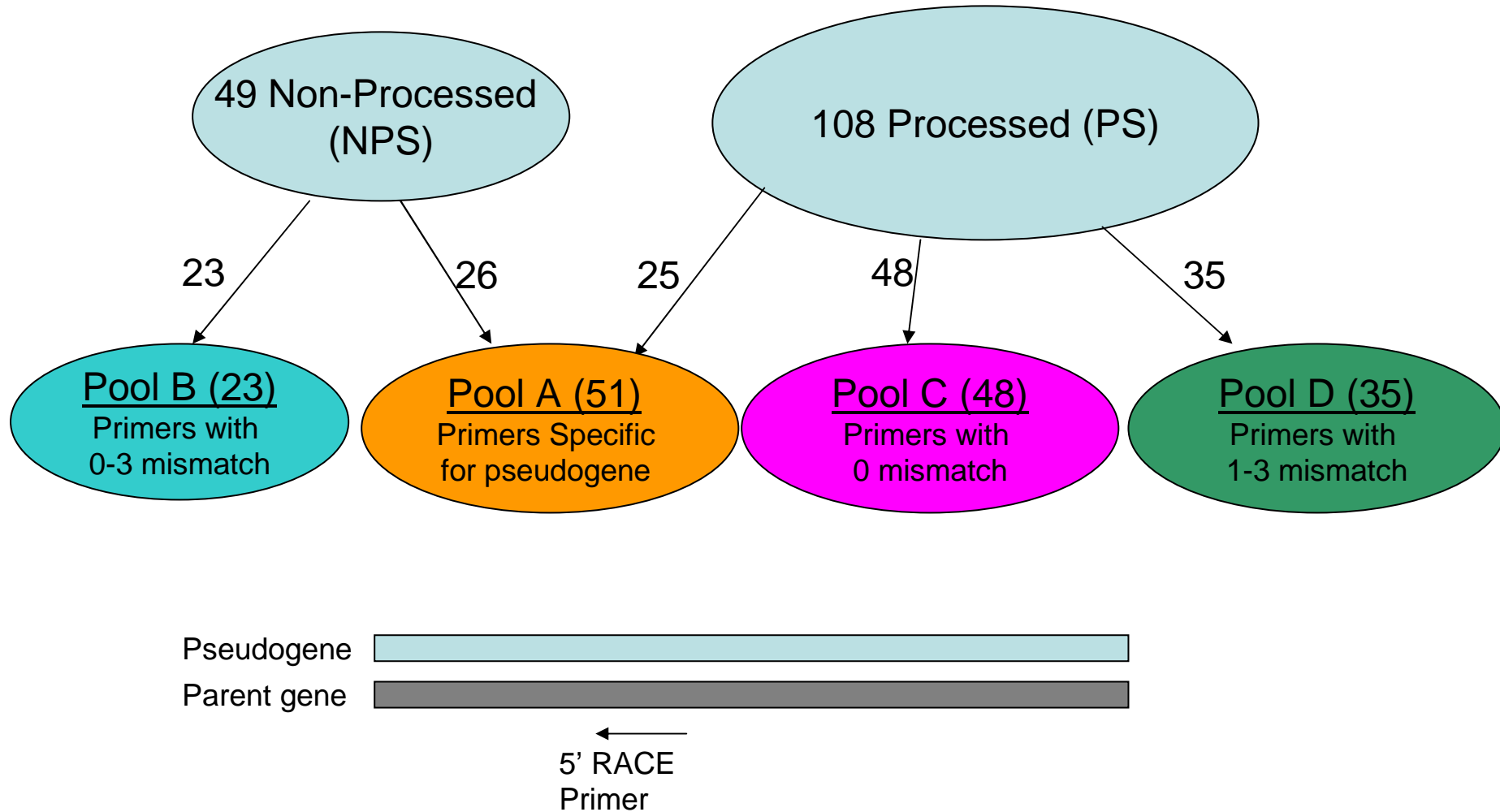
Transcription of ENCODE Pseudogenes -- RACE

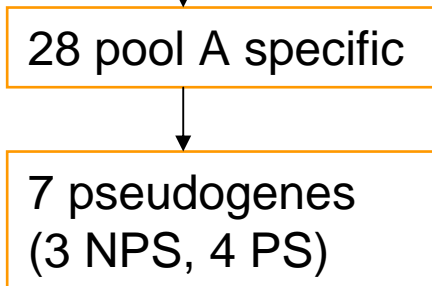
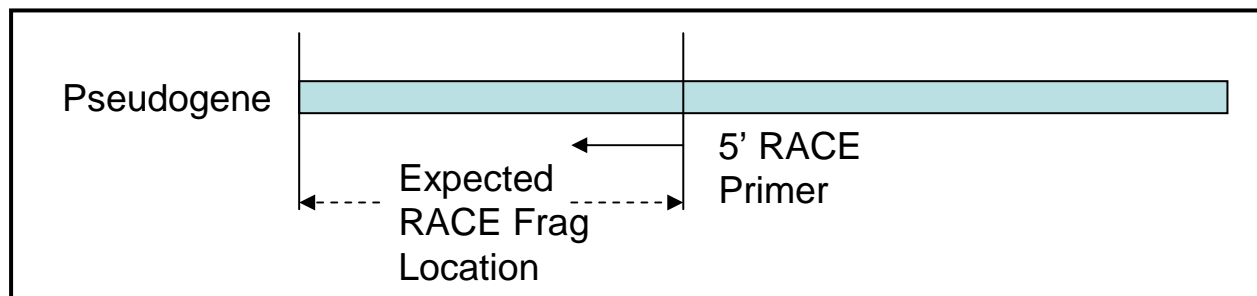
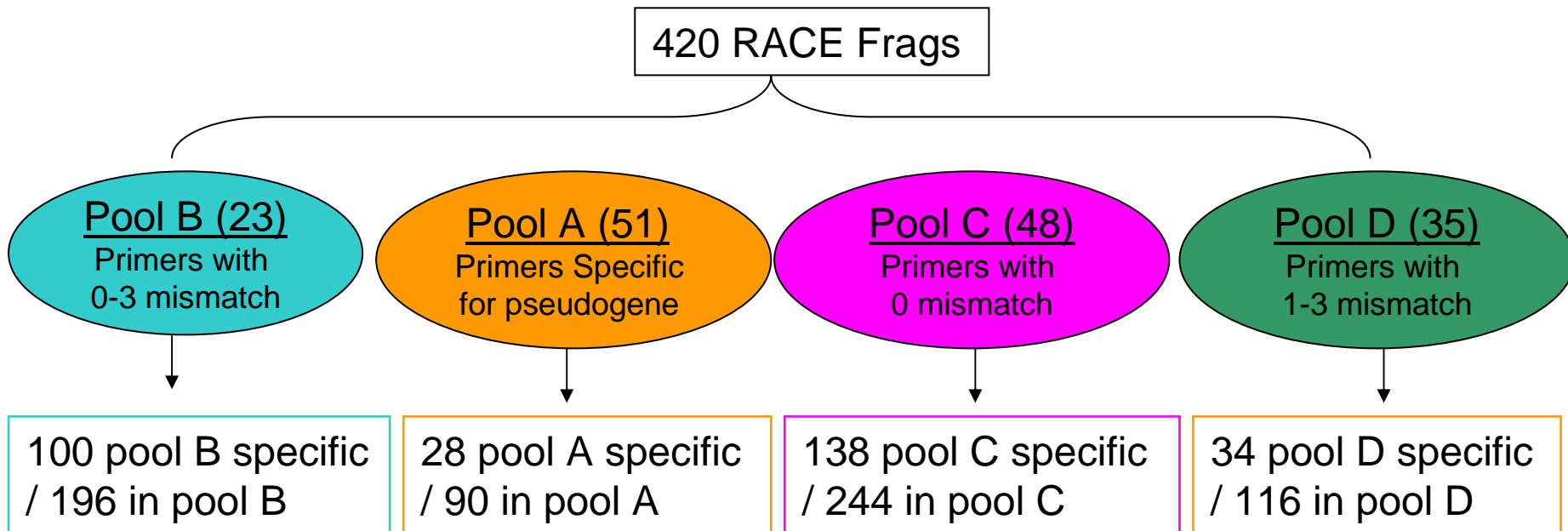
ENCODE Pseudogene Team

Prepared by Deyou Zheng

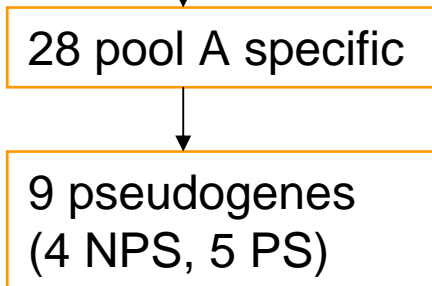
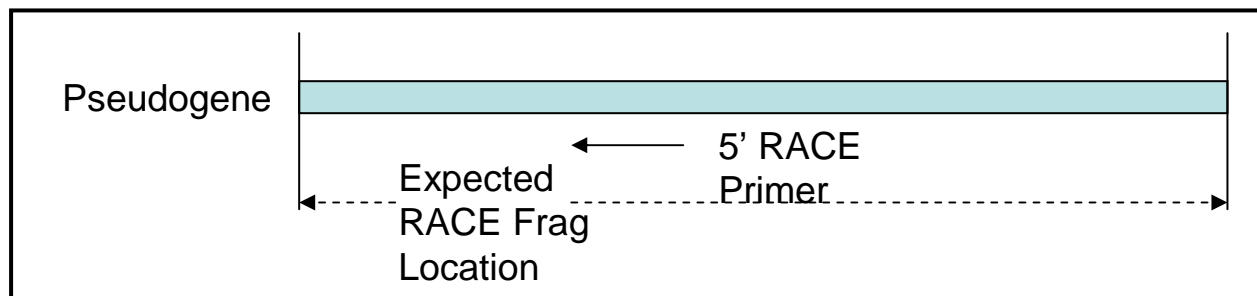
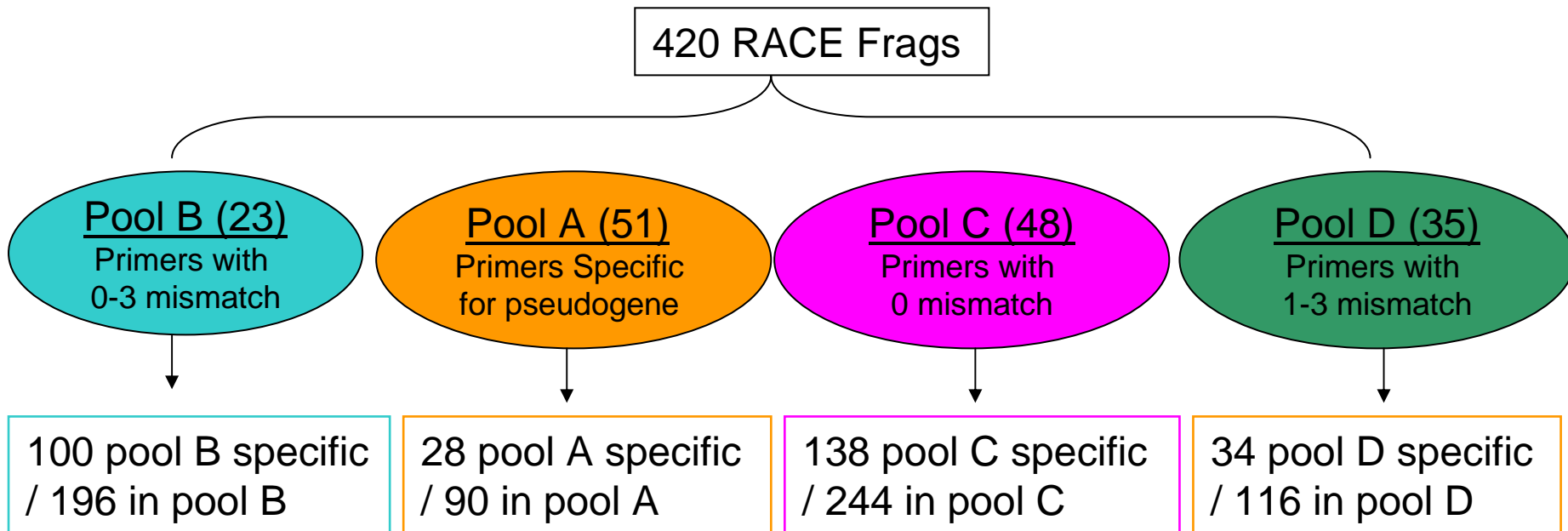
02-09-06

5'-RACE Experiment Design



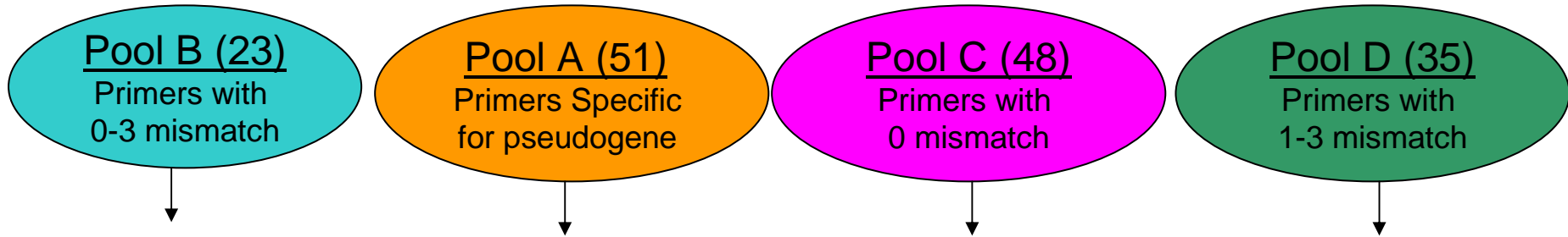


Considered the expected 5' region only, 7 pseudogenes in pool A have RACEFrags.

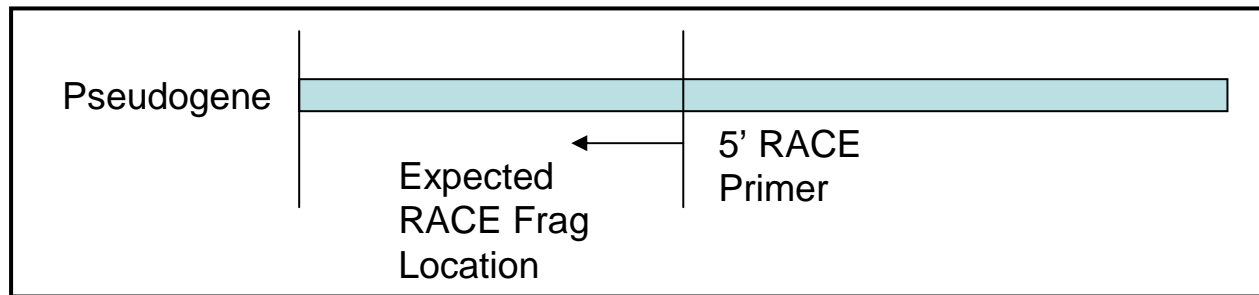


Considered the entire region, 9 pseudogenes in pool A have RACEFrags.

420 RACE Frags



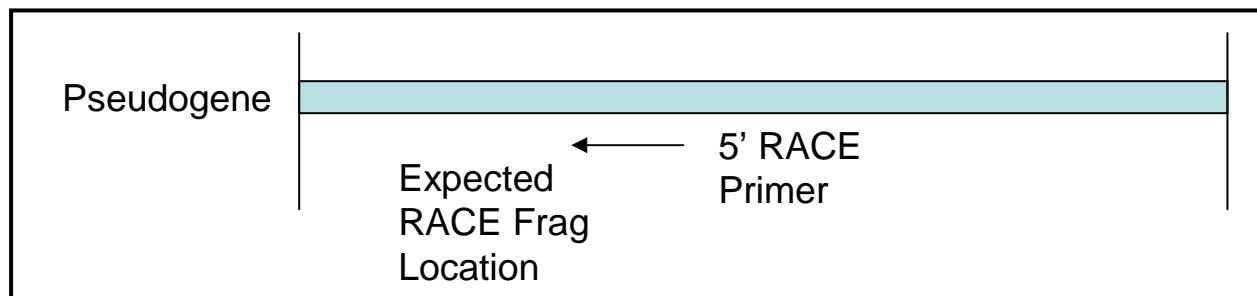
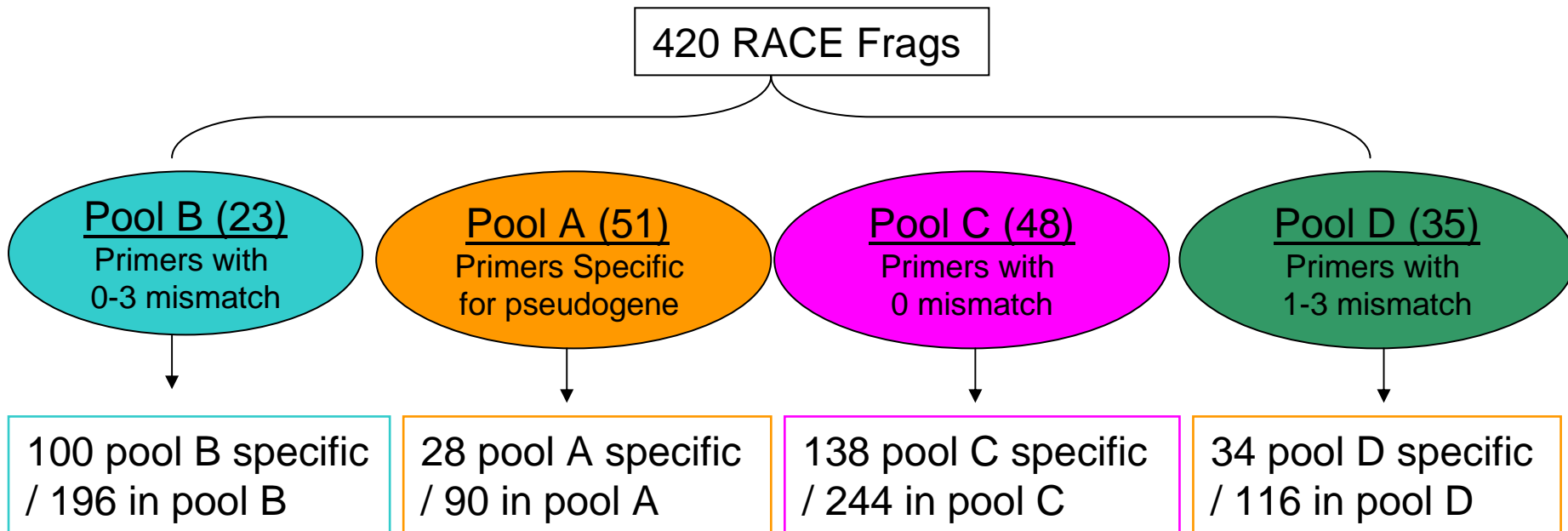
100 pool B specific / 196 in pool B	28 pool A specific / 90 in pool A	138 pool C specific / 244 in pool C	34 pool D specific / 116 in pool D
-------------------------------------	-----------------------------------	-------------------------------------	------------------------------------



100 pool B specific	28 pool A specific	138 pool C specific	34 pool D specific
---------------------	--------------------	---------------------	--------------------

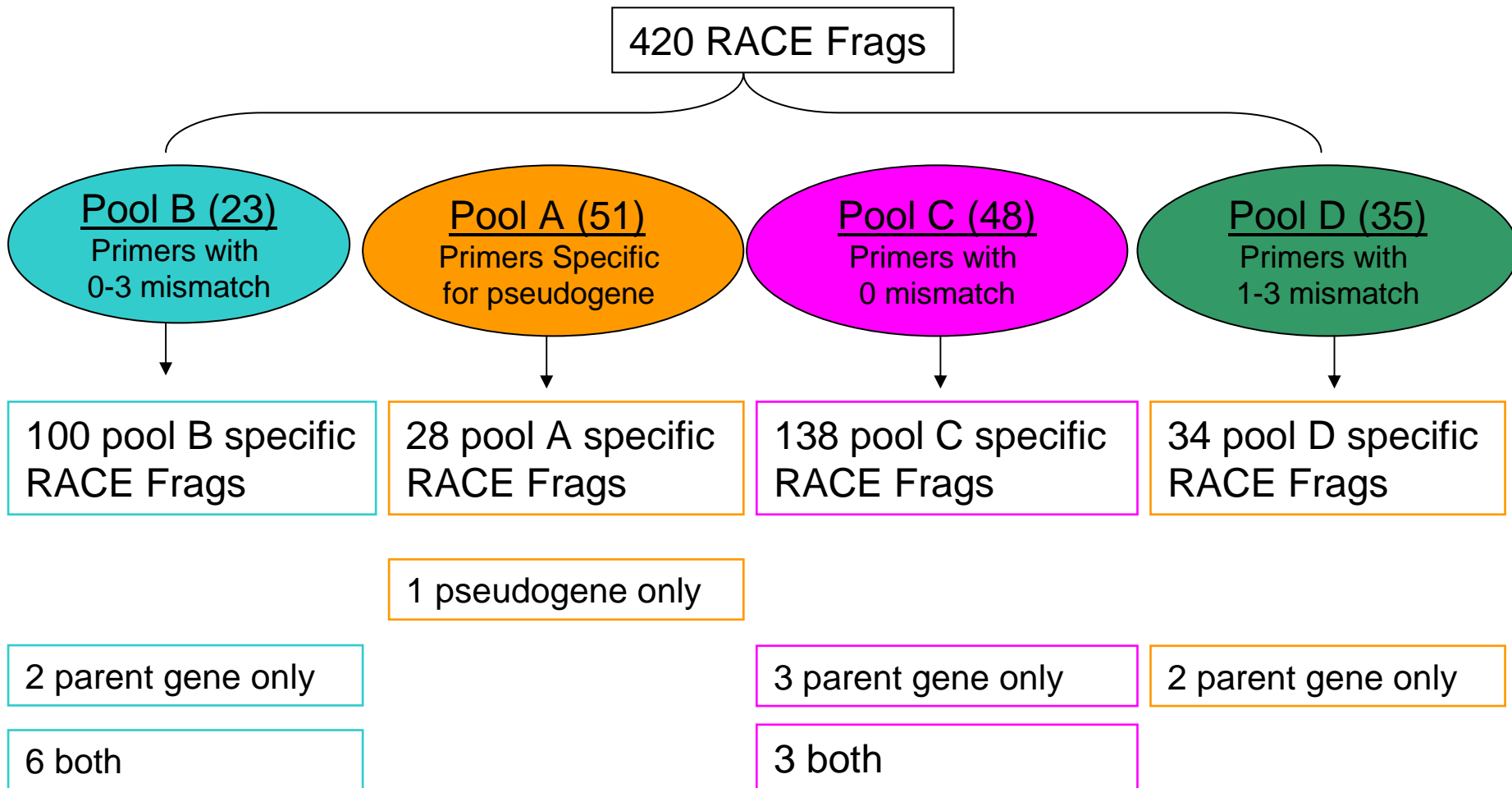
6 pseudogenes (6 NPS)	7 pseudogenes (3 NPS, 4 PS)	24 pseudogenes (24 PS)	7 pseudogenes (7 PS)
-----------------------	-----------------------------	------------------------	----------------------

All pools à 44 pseudogenes have RACEFrags (9 NPS, 35 PS).



100 pool B specific	28 pool A specific	138 pool C specific	34 pool D specific
6 pseudogenes (6 NPS)	9 pseudogenes (4 NPS, 5 PS)	26 pseudogenes (26 PS)	7 pseudogenes (7 PS)

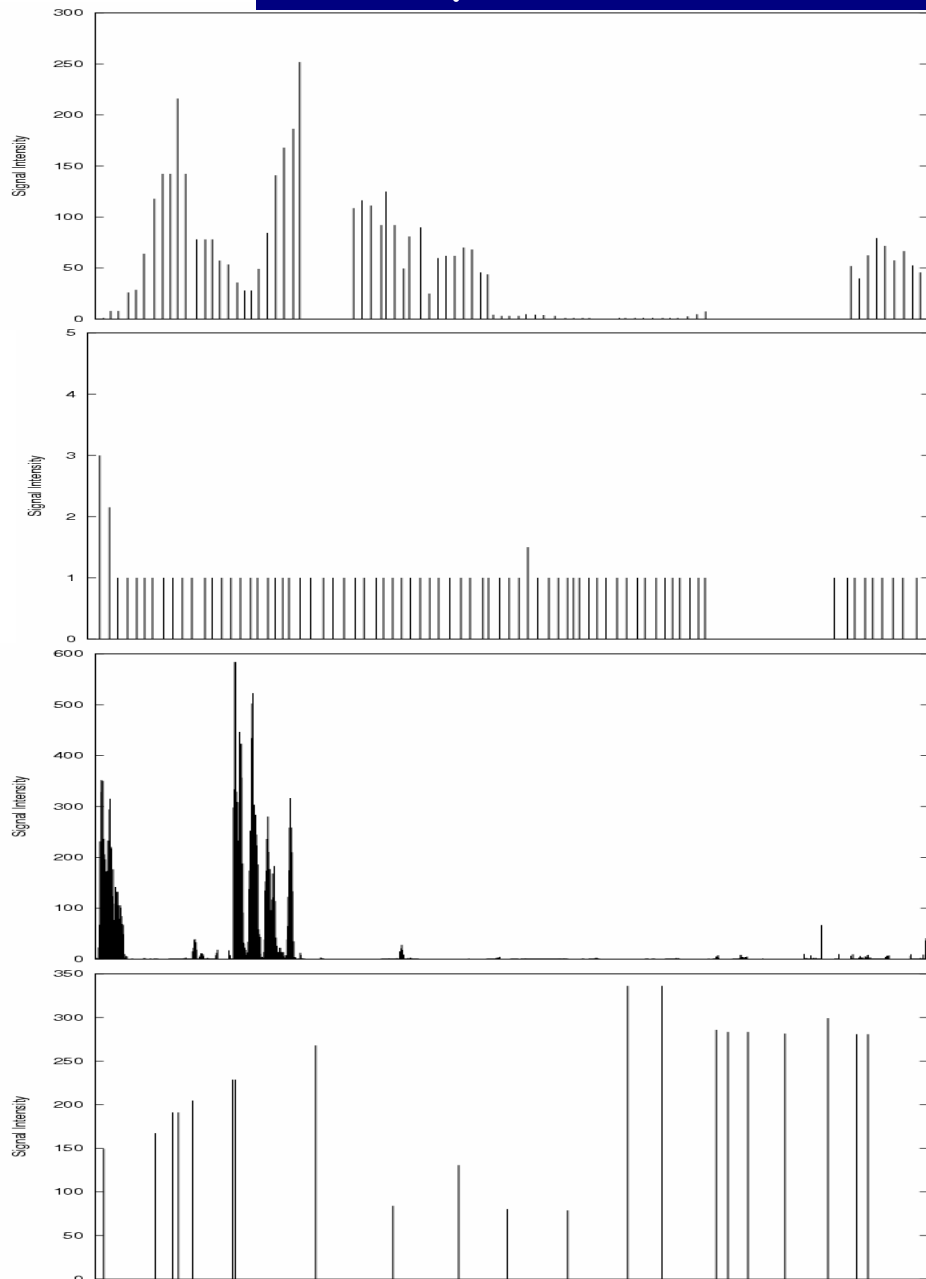
All pools à 48 Pseudogenes have RACEFrags (10 NPS, 38 PS).



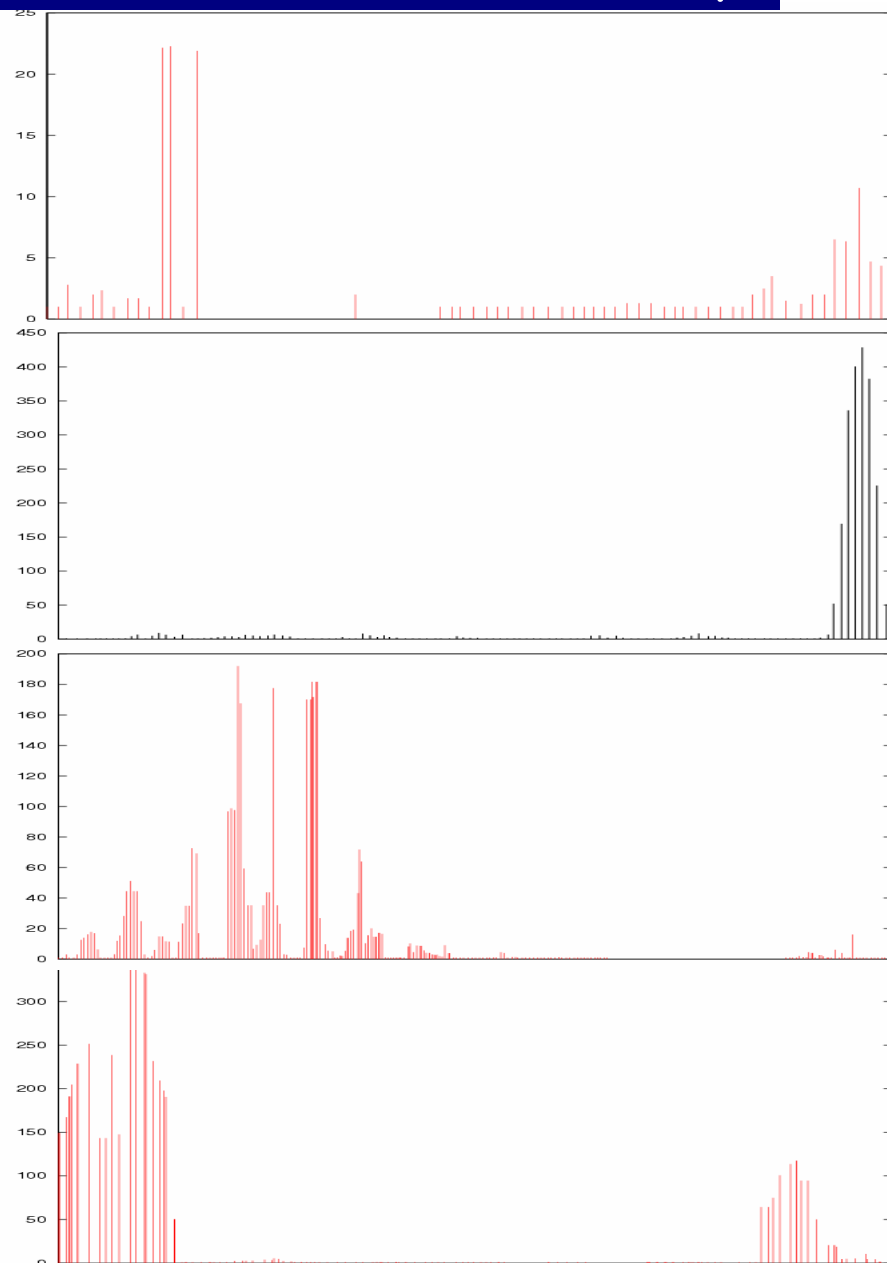
When pseudogenes and parent genes are within ENCODE regions:

- 1 -- pseudogenes transcribed, but not RACEFrag from the parent genes.
- 7 -- parent genes transcribed, but not RACEFrag from the pseudogenes.
- 9 -- both pseudogenes and their parents are transcribed.

Examples of Probe Intensities from RACE-Chip



Pseudogenes



Parent genes

Pseudogenes and Transcription Evidence

	TAR / transfrag	CAGE	DiTag	racefrag	EST / mRNA
TAR / transfrag	105	14	2	2	14
CAGE		14	1	0	1
DiTag			2	0	0
racefrag				8	1
EST / mRNA					21

Methods for intersecting a pseudogene with the five data sets:

