

Perlegen genotyping process and SNPs

GAIN kick-off and analysis workshop

29 November, 2006

David Hinds, Ph.D.

Dennis G. Ballinger, Ph.D.

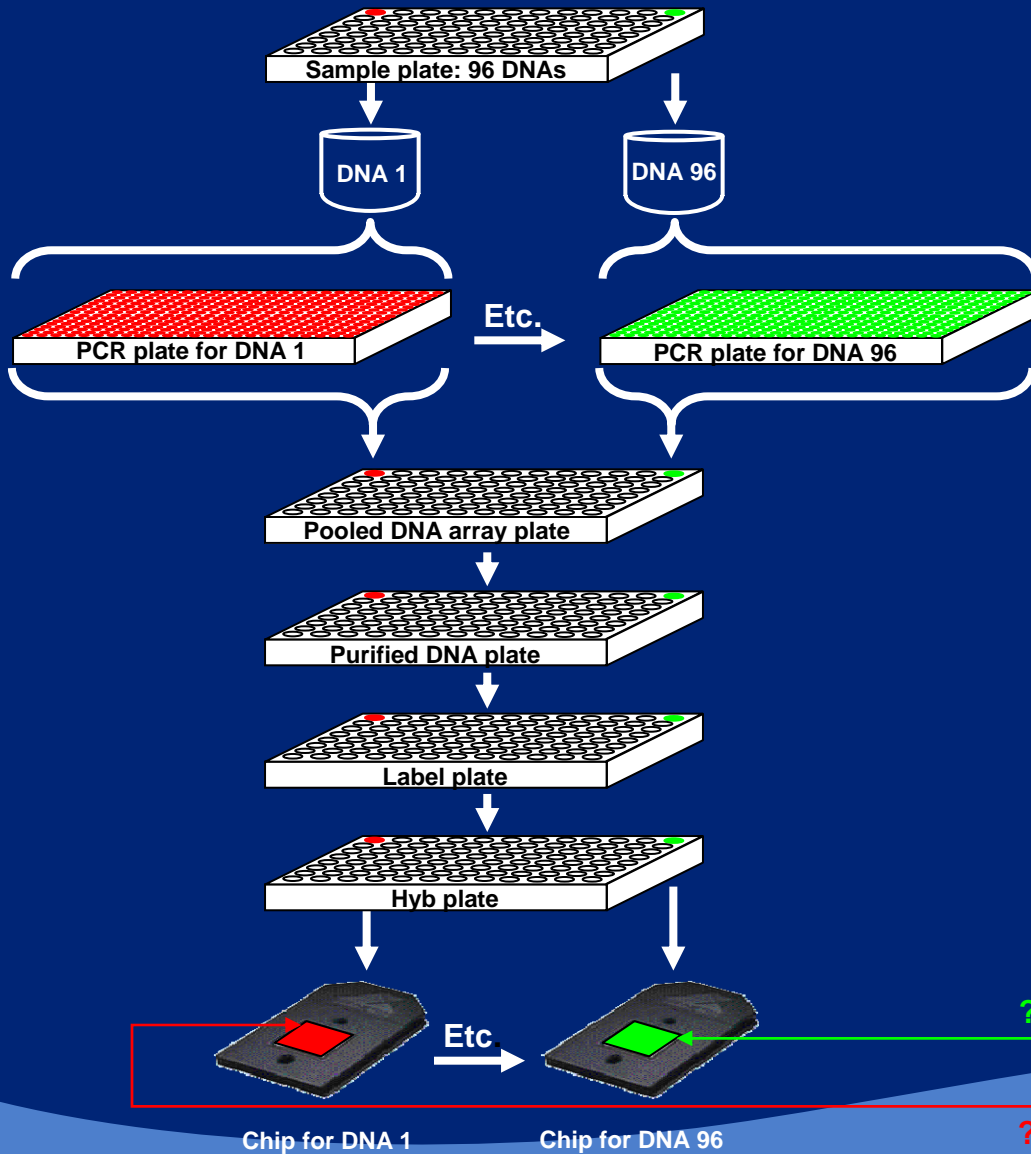
Perlegen Sciences Inc.

dballinger@perlegen.com



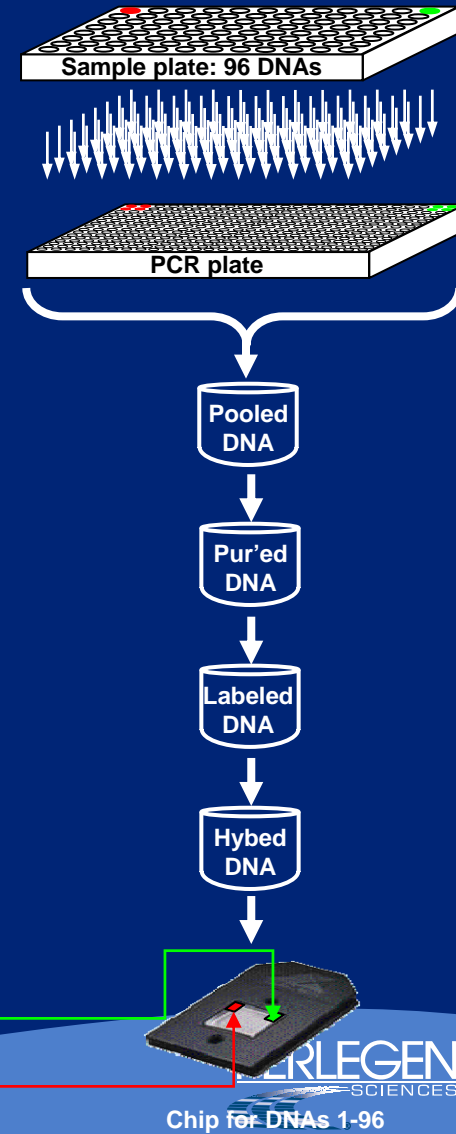
Experimental Genotyping

One sample per chip
All wells in PCR plate have different primers
394 wells of PCR per sample



IDQC Genotyping

96 samples per chip
All wells in PCR plate have different primers
4 wells of PCR per sample



Genotyping

- Adaptive clustering algorithm
- Background corrected ref and alt allele intensities
 - ♦ Could be made available to NCBI in addition to .cel files
- Cluster all samples for a study together

Whole-Genome Patterns of Common Human DNA Variation Have Been Characterized



18 February 2005

.....February, 2005



27 October 2005

.....October, 2005

Integrated HapMap I and II LD map characteristics

HapMap Panel	SNPs	Bins	Singletons
CEU	2,235,139	265,489	274,043
JPT+CHB	2,049,667	237,340	237,641
YRI	2,506,835	388,443	688,964

SNP selection criteria – 600k SNPs

- 600,000 SNPs
 - Approximately 80% are expected to yield working assays.
- Tag SNPs were selected for all CEU bins,
 - Redundant tags selected in bins of 20 or more SNPs.
- Tag SNPs selected based on prior Perlegen performance.
- SNPs selected in two rounds
 - replacements selected for failed assays for larger LD.

Additional SNP selection criteria – 600k SNPs

- Most CEU singletons covered.
 - Priority by location in or near genes and then by MAF.
- Tags were also included for JPT+CHB LD bins of >3 SNPs.
 - Most smaller JPT+CHB bins also covered.
 - 40% of the JPT+CHB singletons also covered.
- Also includes almost 20,000 nsSNPs.

YRI additional SNPs – 750k SNPs

- 150,000 SNPs to improve coverage of the YRI LD map.
- Tags for all YRI LD bins of >2 SNPs.
- Redundant tags for the largest bins of 20 or more SNPs.
- Limited coverage of tags for two-SNP bins and singletons.
 - Preferentially selected from exons.

Maximal r^2 correlation of non-assayed HapMap SNPs

Study Design	$[r^2]$	$r^2 \geq 0.5$	$r^2 \geq 0.8$	Power
CEU	0.94	97%	93%	92%
JPT+CHB	0.91	95%	86%	88%
YRI	0.80	82%	66%	72%

Assumes 80% assay success

50 randomly selected 500 kb intervals

Similar in ENCODE regions

PERLEGEN

SCIENCES

