Insights & Perspectives

# The Power of the Sentrix® HumanHap300 Genotyping BeadChip

An interview with Illumina Senior Bioinformatics Scientist Dr. Pauline Ng.

**What's so special about Illumina's HumanHap300 Genotyping BeadChip?**

The HumanHap300 BeadChip is designed for effective whole-genome association studies. Five years ago, we were genotyping a couple of thousand SNPs per individual. With the HumanHap300, we can now genotype over 300,000 SNPs on a single BeadChip. The ability to interrogate so many SNPs simultaneously allows researchers to efficiently perform meaningful whole-genome association studies to identify the genes involved in complex diseases like cancer, obesity and cardiovascular disease. We are entering an exciting new era, where products like the HumanHap300 will help us to study diseases with unsurpassed speed and confidence.

**Why is there so much interest in this product?**

Technology has finally caught up with the science and we can now perform whole-genome association studies with hundreds of thousands of SNPs. Think about that. That's an incredible number of SNPs - we are giving people the ability to generate about 25 - 50 million genotypes a day. That number is orders of magnitude larger than what could have been generated before. In the past five years, Illumina has generated over 500 million genotypes multiplexing at 1000-fold. With Infinium™ technology, we can surpass this number in just 10 days; this is a very exciting advancement.

**How is this different from other whole-genome products?**

The SNP content and data quality are two things we are extremely proud of. Content is really important because there is so much genetic variation out there. So if you are studying a disease, you are really trying to find just the handful of variations that play a role in the disease in a mix of millions of markers. It's like finding a couple of needles in a haystack, so markers have to be selected wisely. You can pick your SNPs randomly, and hope that by chance your SNP is near the disease SNP, or in genetics terminology, in linkage disequilibrium with the disease SNP. Or you can utilize our current knowledge of the human population and pick representative or "tag" SNPs that maximize your chance for detecting the disease SNP. We have done the latter. The combination of this tagSNP approach with our outstanding data quality ensures the highest quality output.

**What's the strategy for choosing SNPs for the HumanHap300 and how did you come up with it?**

We chose tagSNPs to capture the common variation in HapMap Phase I data. We've tagged genes and conserved regions at a higher density than the rest of the genome. The strategy was developed with an External Content Committee, a committee of leaders in the genetics field from around the world. A lot of different strategies were discussed,

ABOUT PAULINE

Pauline Ng was a key contributor to the content selection for the Sentrix HumanHap300 Genotyping BeadChip. Dr. Ng received a B.S. in Biology from California Institute of Technology, Pasadena, CA and a Ph.D. in Bioengineering from University of Washington- Fred Hutchinson Cancer Research Center in Seattle, WA.

and this is the one everyone thought was best. We're really pleased with the results because we think the content will be incredibly useful to the genetics community.

### Why tagSNPs?

The HapMap Project has provided us with the patterns of human genetic variation, basically an extensive resource that built the foundation to enable the study of genetic links to diseases. By using this information, we can intelligently choose tagSNPs. In the HumanHap300 BeadChip, we captured 80% of the common SNPs in individuals with European ancestry from the HapMap Project at an $r^2 \geq 0.8$ threshold. Without the HapMap Project we would still be choosing random SNPs rather than choosing the most informative SNPs. The HapMap Project made it possible for us to reduce the over one million SNPs to an informative set of more than 300,000. This has huge implications in terms of reducing cost and false positive rate with fewer and more powerful data points.

### What performance factors are important for whole-genome association studies?

Studies are much harder to do with missing data. Generally, these complex disease traits have relatively small gene effects. If the disease SNP is in linkage disequilibrium with a SNP on the HumanHap300, and the call rate for that SNP is low, you may miss the association. This is why we're so proud of our stellar call rate of greater than 99%. The SNPs on the HumanHap300 BeadChip have undergone a rigorous screening process consisting of a bioinformatics filter and two functional screenings before we finally placed it on our HumanHap300 BeadChip. Besides a high call rate, we have outstanding reproducibility and concordance with genotypes obtained from the HapMap project.

ADDITIONAL INFORMATION
To learn more about the HumanHap300 Genotyping BeadChip, visit our website or contact us.

ILLUMINA, INC.
CUSTOMER SOLUTIONS
9885 Towne Centre Drive
San Diego, CA 92121-1975
1.800.809.4566 (toll free)
1.858.202.4566 (outside U.S.)
techsupport@illumina.com
www.illumina.com

FOR RESEARCH USE ONLY


Pub. No. 370-2006-004 01Feb06