illumina®

Omni Whole-Genome DNA Analysis BeadChips

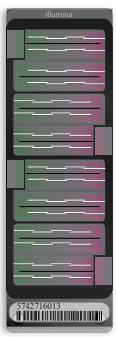
Offering a combination of powerful content and unprecedented flexibility.

Introduction

Whole-genome genotyping (WGGT) microarrays have been used successfully for nearly a decade to identify regions of the human genome that contribute to disease susceptibility and phenotypes. Two primary applications of these tools, genome-wide association studies (GWAS) and identification of copy number variation (CNV), have enabled researchers to achieve a greater understanding of genetic variation and its impact on human health and disease.

In just a few short years, the research community has identified thousands of disease-associated loci and published hundreds of peer-reviewed papers using WGGT technology. Researchers pinpoint causative alleles by genotyping a cohort of DNA samples, then investigating allele frequency shifts between affected and unaffected individuals. The power to detect these associated alleles has steadily grown with the addition of new—and increasingly larger—data sets gathered from next-generation sequencing studies. By leveraging this content, the Omni family of microarrays provides unprecedented coverage of the human genome, empowering researchers to evaluate variants with more powerful technology than ever before. Much of the content was selected from the 1000 Genomes Project (1kGP), a global sequencing effort by research institutions seeking to identify human genetic variants occurring at significant frequencies across diverse populations.

Figure 1: Whole Genome DNA BeadChips

OmniExpress-24	Omni2.5-8	Omni5
		
Highest throughput, exceptional price, common variation coverage down to 5% MAF.	Comprehensive common and rare variation coverage down to 2.5% MAF from the 1kGP.	Near complete common and rare variation coverage down to 1% MAF from the 1kGP.

Omni arrays provide flexibility for timing and budget to help investigators achieve their research goals.

Omni Family of Microarrays

The Omni family of microarrays offers a series of complementary and additive array options for a range of study objectives (Figure 1). Each primary array features a unique set of markers designed to target a specific minor allele frequency (MAF) range (Table 1).

From the flagship HumanOmni5 (Omni5) array targeting variants down to 1% MAF, to the 24-sample HumanOmniExpress (OmniExpress) array, the Omni family of microarrays provides the flexibility to meet a variety of research goals and budgets. Semi-custom options allow researchers to increase the power of their studies by tailoring the arrays with novel variants from their own sequencing data. For follow-up studies, fully custom iSelect® BeadChips can be easily developed with up to 1 million markers targeting any loci across the genome.

The Omni5

The Omni5 BeadChip delivers the complete set of Omni family markers on one array, plus the ability to add 500K custom markers. Content includes powerful tagSNPs selected from the International HapMap and 1000 Genomes Projects that target genetic variation down to 1% MAF. The Omni5 array also includes focused, high-value content covering important regions of the genome such as the MHC, ADME genes, and nsSNPs. The option to add 500K custom markers allows researchers to tailor the BeadChip for targeted applications and population-focused or disease-specific studies. With an average marker spacing of only 680 base pairs, the Omni5 is the industry leader for genotyping and CNV detection.

The Omni2.5

The HumanOmni2.5-8 BeadChip (Omni2.5) features ~2.3 million markers that capture genomic variation down to 2.5% MAF. Optimized tag SNP content and dense marker spacing (mean spacing = 1.2 kb) enable a broad range of study types, including CNV and other structural variation applications. Customized content up to 200K markers can be added to the Omni2.5+ version of the BeadChip.

The OmniExpress

The HumanOmniExpress BeadChip (OmniExpress) delivers excellent power for common-variant GWAS, providing high sample throughput at the industry's best price. This 24-sample BeadChip is the ideal solution for processing the greatest number of samples within a given budget. Optimized tag SNP content from all three phases of the HapMap project has been strategically selected to capture the greatest amount of common SNP variation (> 5% MAF). For researchers needing a more customized solution, up to 50K markers can be added with OmniExpress+ BeadChip. This option provides the same base content as the OmniExpress BeadChips, but allows researchers to include selected markers for their specific study.

AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
TCAACGTACCGTAACGAACGTATCATTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTT
CGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACC
AACGTACCATTAAGAGCTACCGTGCAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAA
AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
GATTACTTGATCCACTGATTCAACGTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAGCAACGAGC
CGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGAAAAGAATGATAACAGTA

Table 1: Omni BeadChip Performance Parameters

	OmniExpress	Omni2.5	Omni5
Number of Fixed Markers	741,000	2,379,855	4,301,331
Available Custom Markers	up to 50,000	up to 200,000	up to 500,000
Number of Samples	24	8	4
DNA Requirement	200 ng	200 ng	400 ng
Assay	Infinium HTS	Infinium LCG	Infinium LCG
Instrument Support	HiScan or iScan	HiScan or iScan	HiScan or iScan
Sample Throughput*	> 2800 / week	~1,067 samples / week	> 460 samples / week
Scan Time / Sample	2.5 minutes	6.5 minutes (HiScan) 11.4 minutes (iScan)	15 minutes (HiScan) 25 minutes (iScan)

% Variation Captured (r^2 > 0.8)	1kGP† MAF > 5%	1kGP† MAF > 1%	1kGP† MAF > 5%	1kGP† MAF > 1%	1kGP† MAF > 5%	1kGP† MAF > 1%
CEU	0.73	0.58	0.83	0.73	0.87	0.83
CHB + JPT	0.74	0.62	0.83	0.73	0.85	0.76
YRI	0.40	0.25	0.65	0.51	0.71	0.58

Data Performance	Value¨ / Product Specification		
Call Rate (average)	99.84% / > 99%	99.65% / > 99%	99.9% / > 99%
Reproducibility	99.99% / > 99.9%	99.99% / > 99.9%	99.99% / > 99.9%
Log R Dev	0.15 / < 0.30‡	0.12 / < 0.30‡	0.12 / < 0.30**

Spacing	Mean / Median / 90th%		
Spacing (kb)	4.0 / 2.1 / 9.3	1.19 / 0.64 / 2.76	0.68 / 0.36 / 1.57

Marker Categories	Number of Markers		
Number of SNPs with 10kb of RefSeq genes	392,197	1,231,382	2,311,849
Nonsynonymous SNPs (NCBI annotated)	15,062	41,900	84,004
MHC / ADME	7,459 / 16,649	19,238 / 27,335	43,904 / 43,615
Sex Chromosome (X / Y / PAR Loci)	18,055 / 1,409 / 471	55,208 / 2,561 / 418	113,213 / 2,498 / 511
Mitochondrial	0	256	267

* Estimate assumes one iScan system, one AutoLoader2, one Tecan Robot, and a five-day work week.

† Compared against June 2011 1kGP data release.

¨ Values are derived from reference samples.

‡ Value expected for typical projects, excluding tumor samples or any samples prepared not following standard Illumina protocols.

Proven Technology

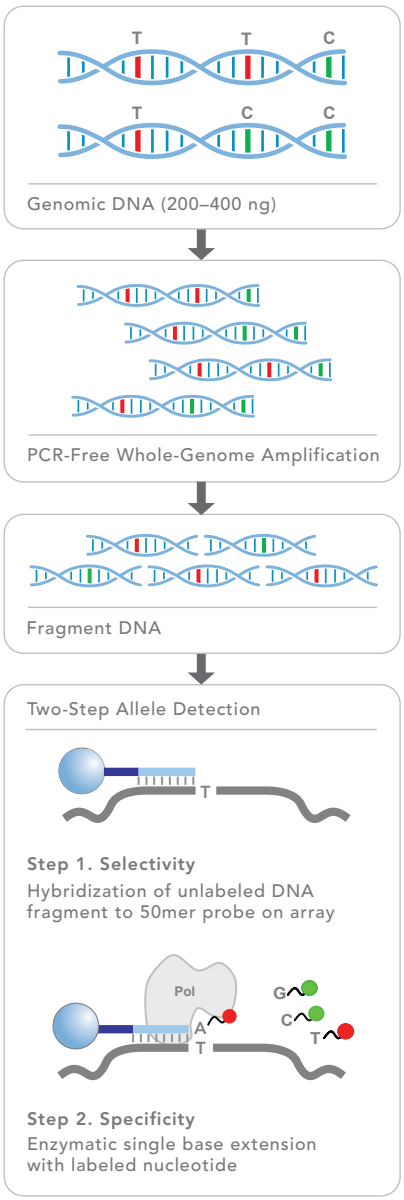
The combination of the proven BeadArray platform, advanced assay technology, and proprietary algorithms present a powerful solution for genetic analysis.

Intelligent tag SNP Content

Illumina's proven tag SNP approach selects the most informative markers from the 1kGP data set for BeadChip content. The power of a tag SNP approach stems from the inherent correlation among markers—the selection of one highly correlated marker serves as a proxy for a number of additional, highly correlated markers across the genome.

The relationship between markers is commonly measured by the correlation coefficient, r^2. A large r^2 value between two markers indicates that they are highly correlated, making them good proxies for each other. At a maximum $r^2 = 1$, two markers are in perfect Linkage Disequilibrium (LD) and can serve as exact proxies for each other. For these markers, only one SNP must be genotyped to know the genotype of the other with high confidence. Illumina DNA analysis products offer unparalleled genomic coverage by leveraging the tag SNP approach. Illumina BeadChips use the highest average r^2

Figure 2: Infinium HD Assay Chemistry



values in the industry, which maximize the likelihood of finding true associations for a given phenotype. By strategically selecting the most powerful tag SNPs, Illumina scientists can ensure maximum power to identify associations, while reducing the redundant information on each BeadChip.

Assay Chemistry

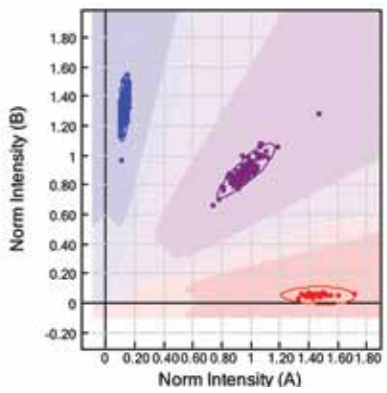
The Infinium Assay has a simple, streamlined workflow (Figure 2). This workflow, as well as the data acquisition and analysis workflows, are common across all products. The Infinium Assay protocol features plate-based sample preparation and whole-genome amplification without PCR or ligation steps, significantly reducing labor and sample handling errors. Following sample DNA hybridization to the BeadChip, two-step allele detection provides high call rates and accuracy. Selectivity and specificity are accomplished in two steps. Target hybridization to bead-bound 50-mer oligos ensures high selectivity while enzymatic single-base extension provides powerful specificity. The single-base extension also incorporates a labeled nucleotide for assay readout. The staining reagent is optimized to provide a higher signal, and more balanced intensities between red and green channels. These features contribute to industry-leading accuracy, high call rates, and copy number data with lower background signal.

Genotype Calling

The Infinium Assay produces two-color readouts (one color per allele) for each SNP in a genotyping study. Intensity values for each color channel, A and B, convey information about the allelic ratio at a single genomic locus (Figure 3). Typical studies incorporate values for large numbers of samples (hundreds to tens of thousands) to ensure significant statistical representation. When these values are appropriately normalized and plotted, distinct patterns (or clusters) emerge, in which samples that have identical genotypes at a given locus exhibit similar signal profiles (A and B values) and aggregate into clusters. For diploid organisms, bi-allelic loci are expected to exhibit three clusters (AA, AB, and BB).

Genotype calls are based on information derived from a standard cluster file, which provides statistical data from a representative sample set. This enables genotypes to be called by referencing assay signal intensities against known data for a given locus. Because the call accuracy is tied to the quality of the cluster data, having an efficient and robust clustering algorithm is essential for accurate genotyping. The proven Illumina Gentrain2 algorithm accurately and efficiently identifies cluster patterns of genotyping samples and reports summary statistics. These statistics are used for downstream genotype calling and CNV analysis. The Gentrain2 algorithm is further described at www.illumina.com/documents/products/technotes/technote_gentrain2.pdf.

Figure 3: Typical Diploid Genotyping Plot



Three clusters of points can be seen for this example locus. The red, purple, and blue regions represent the AA, AB, and BB clusters, respectively.

illumina®

AAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAA
AATCAACGTACCGTAACGAACGTATCATTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACC
AACGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTA
TTAACGTACCATTAAGAGCTACCGTGCAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGA
AAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAA
AAGATTACTTGATCCACTGATTCAACGTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAGCAACGA
AACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGCAACGAAAAGAATGATAACAC

BeadArray Manufacturing

Illumina BeadArray technology is based on small silica beads that self assemble in microwells on planar silica slides. Each bead is covered with thousands of specific oligonucleotide copies that act as the capture sequences in the Infinium Assay. Following self-assembly, a proprietary decoding process maps the location of every bead, ensuring that each one is individually identified and quality controlled. This manufacturing process ensures that, not only each BeadChip, but *each bead* on the array undergoes rigorous testing and quality control.

Exceptional Data Quality

The Omni family of microarrays is powered by the Infinium Assay—the industry's most trusted, proven DNA analysis platform for genotyping and CNV studies. The assay utilizes proprietary Illumina BeadArray™ technology, which allows Omni arrays to deliver a high degree of flexibility, including multiple sample formats and a wide range of SNP assays per array. Infinium products deliver exceptionally high-quality data with respect to call rates (average > 99%), reproducibility (> 99.9%), and low sample redo rates (Table 1). High signal-to-noise ratios and low overall noise levels allow for precise, reliable copy number analysis. With such high data quality, the assay minimizes the number of false positives, allowing researchers to avoid time-consuming additional analysis and expensive follow-up studies on erroneous associations.

Maximized Genomic Coverage

Genomic coverage is a key metric for any whole-genome microarray; it indicates the percent of variation captured on the array at an LD of $r^2 > 0.8$. Before the 1kGP, genomic coverage statistics were based on a catalog of variants identified by the International HapMap Project. This data set contained ~3.5 million variants, targeting MAFs > 5%. In light of the more comprehensive data available from the 1000 Genomes Project, the reference point for coverage statistics have changed. As shown in Table 1 and Figure 4, Omni arrays offer greater than 80% coverage of variants, with respect to 1kGP data, targeting (MAF > 1%).

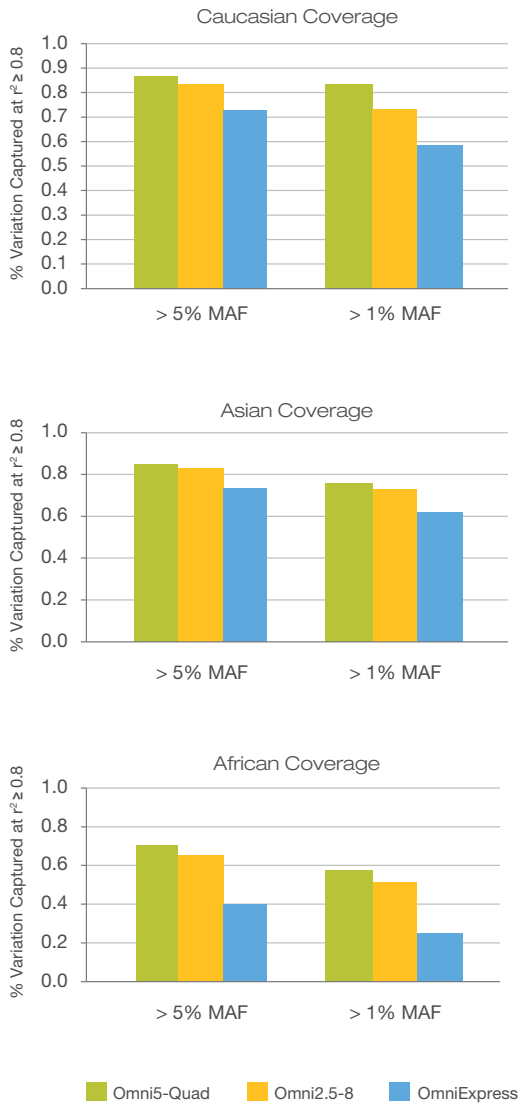
Structural Variation Analysis

Structural variation is thought to be a significant contributor to the genetic basis of human disease. Dense genome-wide coverage on Omni microarrays, coupled with the sensitive Infinium Assay, offer researchers a powerful tool for structural variation analysis. The assay delivers high signal-to-noise ratios and low overall background levels, which are ideal for precise structural variation analysis. Whether performing genotype calling, structural variation analysis, or both, Omni arrays provide a single solution for multiple applications.

Customized For Targeted Studies

Custom iSelect BeadChips can be designed to target specific regions of interest, allowing researchers an ideal selection of markers for a range of budgetary and throughput requirements. The Illumina iSelect custom genotyping platform offers all of the benefits of standard Infinium® products, including industry-leading data quality, streamlined workflows, and informed SNP selection. Custom iSelect BeadChips also include the flexibility to access virtually the entire genome.

Figure 4: Maximized Genomic Coverage



Omni whole-genome microarrays provide superior coverage of common and rare variants across Caucasian, Asian, and African populations. The Omni5 array provides the highest coverage of vairants down to 1% MAF. With the option of selecting 500K additional custom variants on the Omni5, researches can use tagSNPs to increase coverage by up to 10%.

Custom products can be deployed on 4,12, or 24-sample formats. Convenient online tools and Illumina representatives are available to help researchers design and select markers that best suit specific research goals. For more information on iSelect Custom BeadChips, visit www.illumina.com/applications/genotyping/custom_genotyping.ilmn.

Ordering Information

Product	Catalog No.	Product	Catalog No.
HumanOmniExpress-24 v1.0 DNA Analysis BeadChip Kits		**HumanOmniExpress-24+ v1.0 DNA Analysis BeadChip Kits**	
48 samples	WG-312-3001	48 sample	WG-312-3005
288 sample	WG-312-3002	288 sample	WG-312-3006
1152 sample	WG-312-3003	1152 sample	WG-312-3007
HumanOmni2.5-8 v1.1 DNA Analysis BeadChip Kits		**HumanOmni2.5-8+ v1.1 DNA Analysis BeadChip Kits**	
16 sample	WG-312-2511	16 sample	WG-312-2521
48 sample	WG-312-2512	48 sample	WG-312-2522
96 sample	WG-312-2513	96 sample	WG-312-2523
384 sample	WG-312-2514	384 sammple	WG-312-2524
HumanOmni5-Quad v1.0 DNA Analysis BeadChip Kits		**HumanOmni5-Quad+ v1.0 DNA Analysis BeadChip Kits**	
16 sample	WG-311-5001	16 sample	WG-311-5005
48 sample	WG-311-5002	48 sample	WG-311-5006
96 sample	WG-311-5003	96 sample	WG-311-5007
384 sample	WG-311-5004	384 sample	WG-311-5008

Data Analysis Software

Illumina GenomeStudio® Data Analysis Software offers integrated genotyping and copy number tools as well as the Genome Viewer for graphic visualization. GenomeStudio has an open plug-in interface to integrate third-party applications for more downstream data analysis options. Beeline Software provides a direct path to project creation and sample management for large array experiments. The time required for data analysis is reduced by flexible allele calling and data filtering prior to entry into GenomeStudio. Learn more about GenomeStudio and Beeline by visiting support.illumina.com/array/array_software.ilmn. The illumina•Connect program leverages this open architecture and has made numerous plug-ins available to support genotyping and copy number analysis. The illumina•Connect plug-ins are further described and available for download at www.illumina.com/software/illumina_connect.ilmn.

BeadArray Scanners and Automation Systems

Omni micorarrays are compatible with the Illumina iScan and HiScan systems. These cutting-edge array scanners feature high-performance lasers and powerful optical systems that enable rapid scan times and precise assay detection. A convenient modular design enables researchers to easily build out the systems for evolving research needs.

An optional Laboratory Information Management System (LIMS) is available to accurately and efficiently track samples. For more information on Illumina LIMS, visit www.illumina.com/software/lims.ilmn. Robotic automation capabilities can be added to improve throughput for labs processing large numbers of samples. More information on robotic automation of Illumina microarrays can be found at www.illumina.com/products/iscan_system_automation_option_packages.ilmn.

Illumina FastTrack Microarray Services

Illumina FastTrack Microarray Services are available to analyze samples in a timely fashion, at a reasonable cost, using any Infinium DNA Analysis BeadChip. This option allows researchers to acquire high-quality data for limited studies or before purchasing their own equipment. For more information on Illumina FastTrack Services, visit www.illumina.com/services.ilmn.

Learn More

To learn more about the Omni Array Family, visit www.illumina.com/applications/genotyping/omni_family.ilmn.

Illumina, Inc. • 5200 Illumina Way, San Diego, CA 92122 USA • 1.800.809.4566 toll-free • 1.858.202.4566 tel • illumina.com

FOR RESEARCH USE ONLY

© 2010–2013 Illumina, Inc. All rights reserved.
Illumina, illumina*Dx*, BaseSpace, BeadArray, BeadXpress, cBot, CSPro, DASL, DesignStudio, Eco, GAIIx, Genetic Energy, Genome Analyzer, GenomeStudio, GoldenGate, HiScan, HiSeq, Infinium, iSelect, MiSeq, Nextera, NuPCR, Sentrix, SeqMonitor, Solexa, TruSeq, TruSight, VeraCode, the pumpkin orange color, and the Genetic Energy streaming bases design are trademarks or registered trademarks of Illumina, Inc. All other brands and names contained herein are the property of their respective owners.
Pub. No. 370-2010-027 Current as of 08 October 2013

illumina®

AAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAA
AATCAACGTACCGTAACGAACGTATCATTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACC
AACGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTA
TTAACGTACCATTAAGAGCTACCGTGCAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGAT
AAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAA
AAGATTACTTGATCCACTGATTCAACGTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAGCAACGA
AACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGAAAAGAATGATAACA