SNPbrowser™ Software v3.5

A Free Software Tool for the Knowledge-Driven Selection of SNP Genotyping Assays

- Easily visualize SNPs integrated with a physical map, linkage disequilibrium maps, and haplotype blocks
- Stand-alone database contains over six million SNPs, including over 650 million genotypes generated for over three million SNPs validated by the International HapMap Project or Applied Biosystems in five major populations
- Easily select and order SNP assays associated with TaqMan® SNP Genotyping and SNPlex™ Genotyping System Assays with a direct link to the Applied Biosystems online store
- Features high value SNP sets including a large collection of non-synonymous coding SNPs and those associated with disease mutations
- Tagging SNP wizard selects maximally informative SNPs based on user selected parameters providing for more affordable study design
- Optimize marker coverage using the SNP density wizard
- Assess statistical power for association studies by sample size and disease minor allele frequency values
- Provides offline access to the HapMap SNP data set from the International HapMap Project

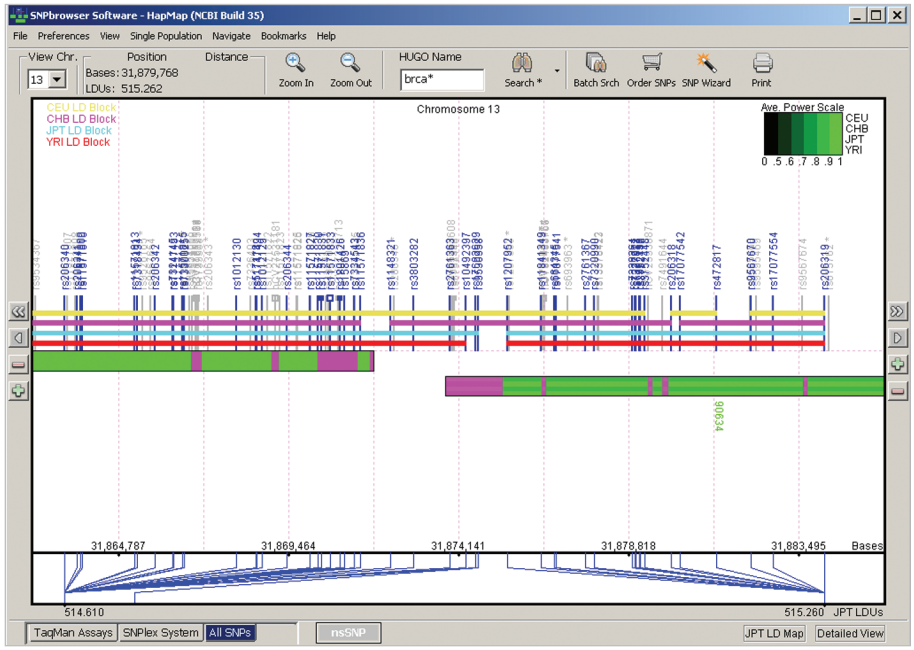

Figure 1. Applied Biosystems SNPbrowser Software v3.5 is a free tool facilitating the easy and intuitive selection of SNPs for successful association and fine mapping studies.

Introduction

Applied Biosystems SNPbrowser™ Software v3.5 is a tool that facilitates the easy and intuitive selection of human SNPs and associated TaqMan® SNP Genotyping Assays. It includes visualization of SNPs integrated with the physical genome maps, linkage disequilibrium (LD) maps, and putative haplotype block information. The LD and haplotype block information is generated from genotypes of over three million SNPs validated by either Applied Biosystems or the International HapMap project in a total of five major populations. This extensive testing resulted in a unique dataset of 650 million genotypes, which can be exported to a file even if unassociated with Applied Biosystems product offerings.

SNPbrowser Software contains a library of over six million human SNPs, facilitating the knowledge-driven design of association and fine mapping studies through the selection of the most informative SNPs for meaningful and powerful study results. The software features assay product associations for over three million TaqMan® SNP Assays and 5.6 million SNPlex System-compatible SNPs available from the Applied Biosystems online store

(**www.appliedbiosystems.com**). The TaqMan SNP Genotyping Assay set features a large selection of TaqMan Validated, Pre-Designed, and non-synonymous Coding SNP Genotyping Assays that are easily identifiable in the SNPbrowser views. The software features the complete high quality HapMap SNP set from the International HapMap SNP project, including over three million SNP assays, offering an offline version of this data that can be supplemented with user specific SNPs of interest.

SNPbrowser Software supports association and fine mapping study designs by offering a wide variety of tools to search, view, export SNP information, and purchase assays for available SNPs. When selecting genomic regions to view, users can easily apply filter criteria such as product associations and SNP minor allele frequency values by populations. The software's Map View also shows gene structure, haplotype blocks, and LD values (Figure 1). Views can be customized and saved as user preferences for future use. The software allows for multiple window views for comparing SNP selection methodologies. SNPbrowser also features SNP wizards that perform back end calculations to optimize and streamline SNP selections in regions of interest. These include a tagging SNP selection wizard and SNP density selection wizard. An extensive Help Text file is provided within the tool to assist in familiarizing users of the extensive functionality offered.

Getting Started is Easy

SNPbrowser Software can be downloaded or ordered on CD for free directly from the Applied Biosystems Web site (**www.allsnps.com/snpbrowser**). Whether you're installing for the first time, or simply upgrading, the install wizard guides you through the installation process to load the database and software tools on your operating system.

Applied Biosystems and HapMap Linkage Disequilibrium Maps

Underlying SNPbrowser Software's visualization tool is a vast amount of data including highly informative metric linkage disequilibrium maps. This information can be used to select a reduced set of highly informative SNPs needed for a genetic association study.

Linkage disequilibrium (LD) refers to the fact that particular alleles at nearby variant sites can co-occur by chance more often than is expected. The existence of linkage disequilibrium means that genetically-linked SNP markers can be used to discover nearby causative mutations that increase susceptibility to disease or adverse drug reactions.

The degree of allelic association is expressed in linkage disequilibrium units (LDUs), which represent distances proportional to the strength of LD. LDUs are analogous to the centi-Morgan units that are commonly used for linkage studies because they provide locations for markers with distances that are additive. These units provide a more appropriate metric than physical distance to select appropriately spaced subsets of SNP markers for genetic association studies.

Two separate databases of validated SNP genotypes, generated for four distinct populations each, are included in SNPbrowser Software. Every time SNPbrowser is launched, the user may select which data set will be used as the backbone information in computing LD maps, putative haplotype blocks, statistical study power, and Tag SNP selections. These two databases are:

Applied Biosystems LD Map Database

During the development of the TaqMan® Validated SNP Genotyping Assays, Applied Biosystems generated the genotype data necessary to empirically create this map. The underlying data was derived from 160,000 SNPs in 45

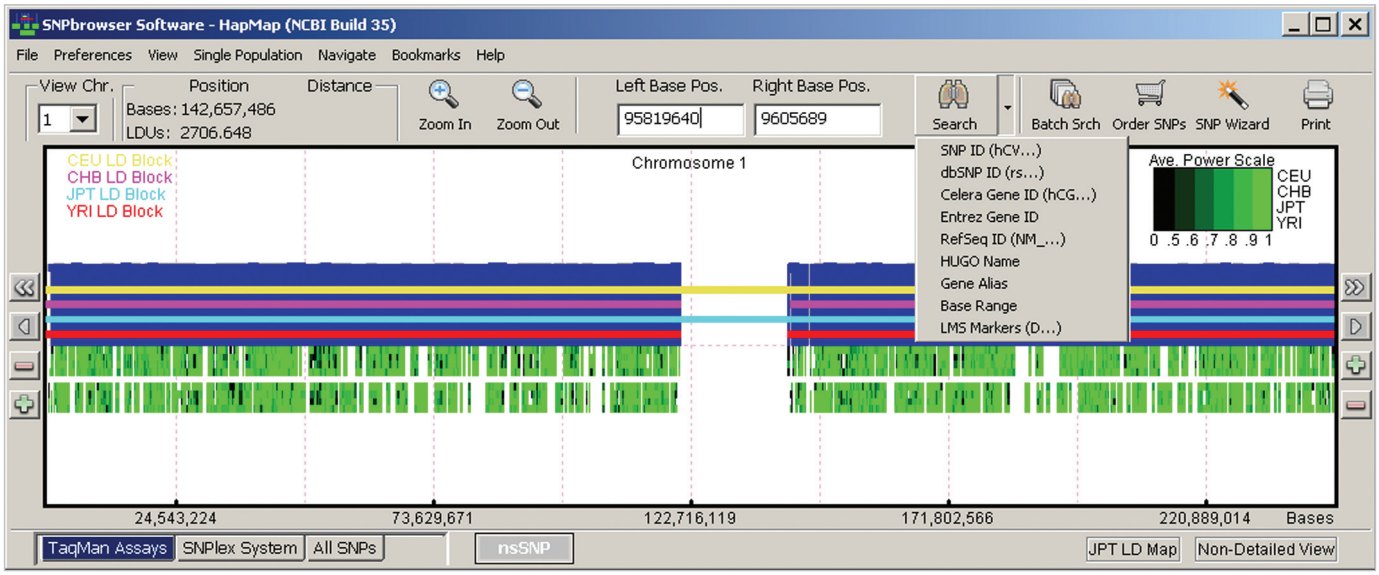

Figure 2. Jump to any chromosome to browse local SNPs, or search using SNP, gene, and marker identifiers.

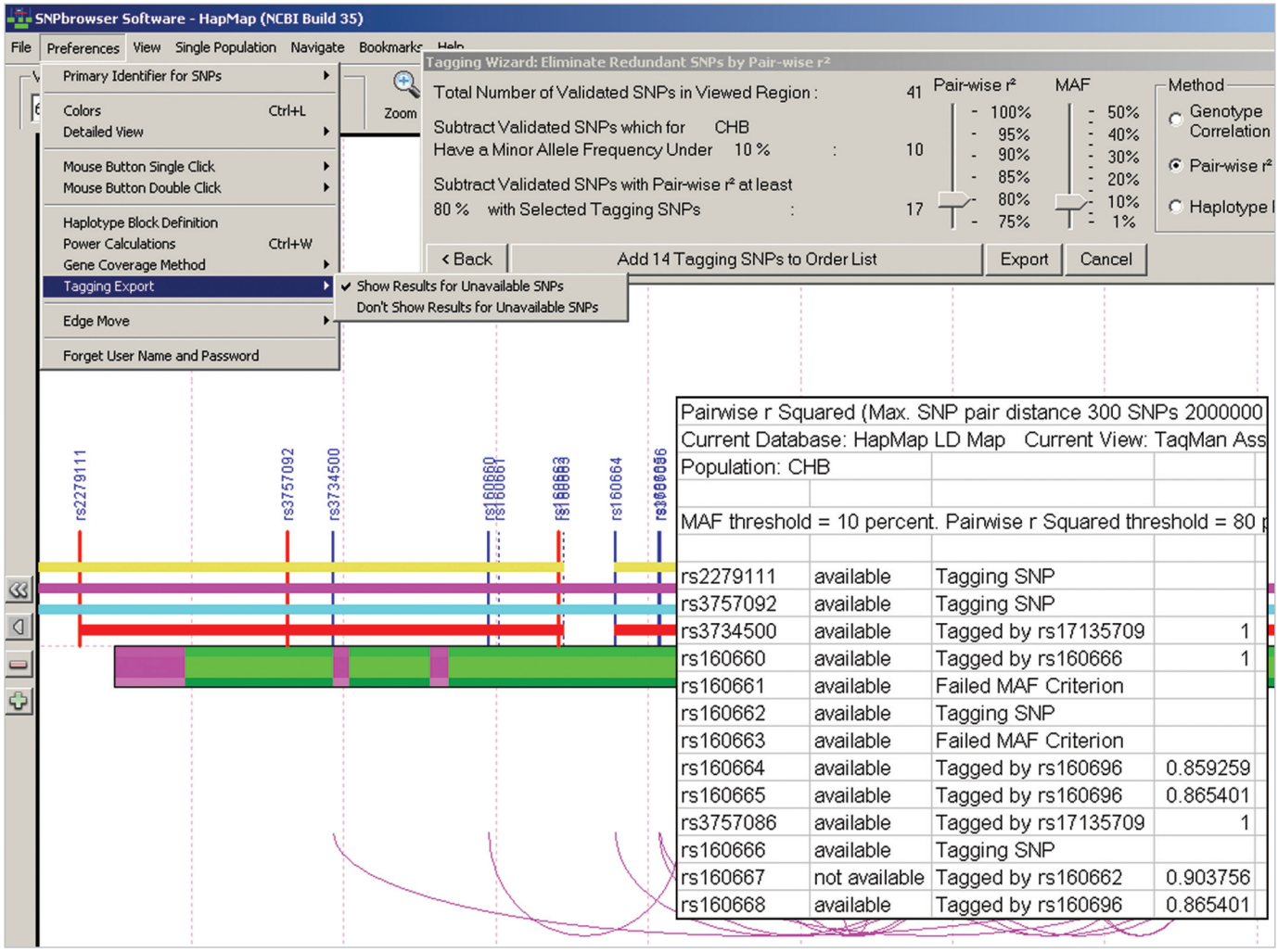

Figure 3. The tagging SNP wizard simplifies the elimination of redundant SNPs with rapid, real time tagging SNP selections. The new advanced export file features allow for export of SNPs with and without product association.

distinct individuals from two to four populations including Caucasian, African American, Japanese and Chinese.

HapMap Project LD Map Database

For this data set, the underlying SNP genotype data was obtained from the International HapMap Project (**www.hapmap.org**) for over 3.2 million high quality SNPs genotyped in 44 to 60 unrelated individuals from four populations including Yorubans, Caucasian, Japanese and Chinese.

Powerful Search and Visualization Tools to Navigate the Genome

The existence of so many SNPs throughout the genome is both an advantage and a disadvantage. While it makes SNPs ideal markers for genetic studies, it can also make finding the right SNPs overwhelming. To solve this problem, SNPbrowser Software gives you a simple, intuitive tool to visualize SNPs throughout the entire genome, and leverage the vast amount of data underlying the TaqMan SNP Genotyping and SNPlex Genotyping System Assays.

Researchers need tools to design studies with high statistical power while minimizing the number of SNP assays necessary for the study. SNPbrowser has been designed to address these challenges to ultimately reduce overall study costs and time to results. SNPbrowser Software v3.5 enables users to locate SNPs of interest by supporting search options with respect to chromosomal regions, genes, or by entering specific search terms including gene identifiers, SNP identifiers or chromosomal locations (Figure 2). A batch search option provides a direct route to locating SNPs of interest. Multiple launches of SNPbrowser may be simultaneously used allowing for a variety of comparative analyses.

Filters can be applied and saved to customize the results views based on specific needs. These include SNP product associations, minor allele frequency values, populations, LD maps, haplotype (LD) blocks, LMS markers, tagging SNP designations, LDU density and disease allele associations.

To support independent researchers, SNPbrowser Software easily enables the import of SNPs from outside sources and incorporates them into the view

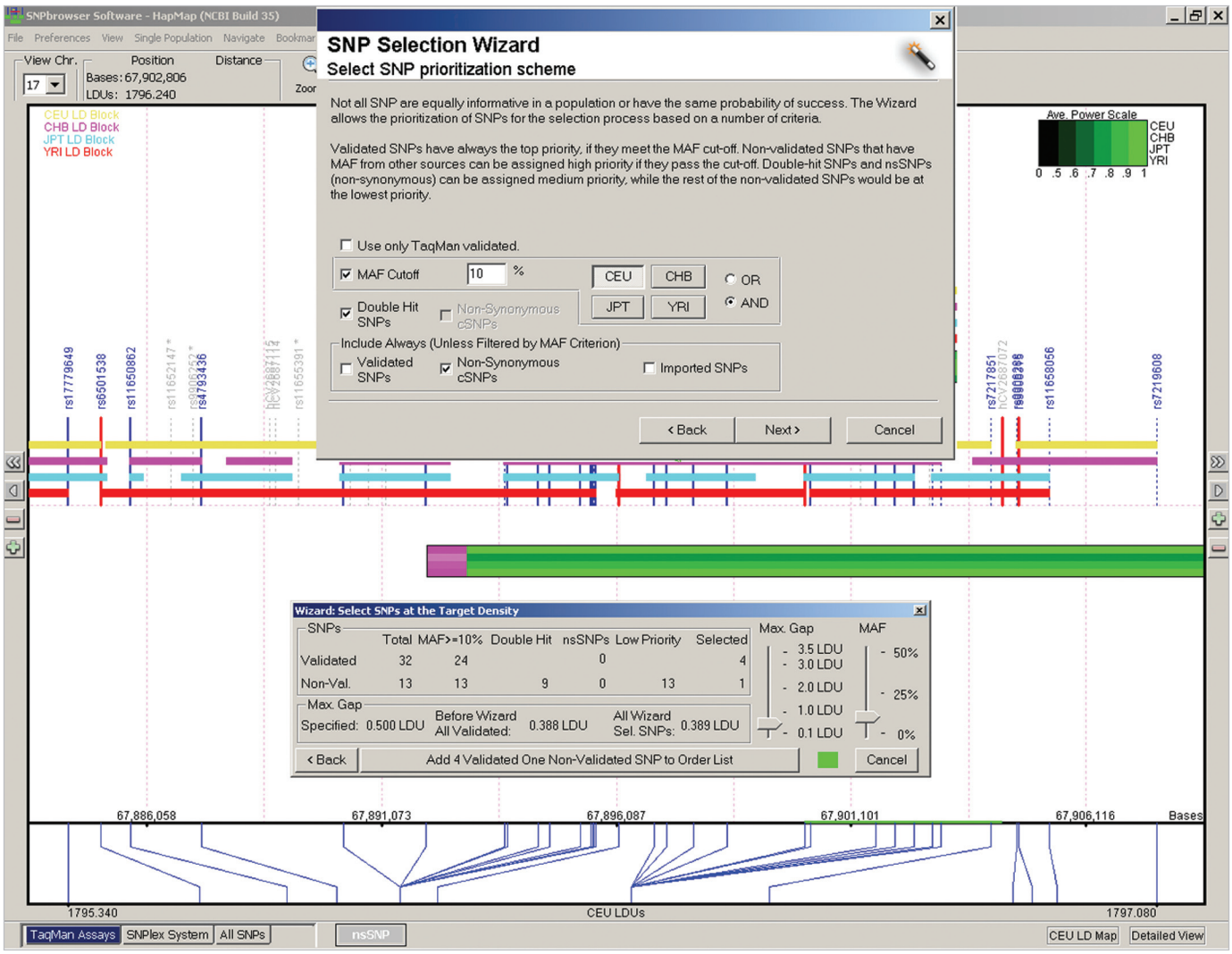

Figure 4. The SNP density wizard allows for a "picket-fence" method to select evenly spaced SNPs using genetic (LD) or physical (kb) distances.

with those in the SNPbrowser database. SNP assays not available through SNPbrowser Software may be ordered as custom assays through either Applied Biosystems Custom TaqMan® SNP Genotyping or custom SNPlex™ Genotyping System Assays services.

Tools for Tagging SNP Selection and Optimizing Marker Coverage

SNPbrowser Software easily enables the selection of tagging SNPs with the inclusion of the tagging SNP selection wizard. Different methods have been proposed to select optimal subsets of SNPs, or "tagging" SNPs, based on the empirical pattern of LD for a given population and genomic region. SNPbrowser Software incorporates three tagging SNP algorithms to facilitate the selection of tagging SNPs and eliminate SNPs providing redundant genetic information. These methods include genotype correlation, pairwise r^2 and haplotype R^2, and aim to retain the statistical power of association achieved for a study while reducing the number of SNPs actually genotyped.

Calculations are performed on-the-fly based on the validated SNPs within a selected genomic region, and tagging SNPs are easily identified on the view (Figure 3). Tagging SNP selection results can be exported to a tab delimited file that can be configured to include all validated SNPs, independent of product associations.

Through the visualization of physical distances, haplotype block boundaries, and linkage disequilibrium units in the SNPbrowser Software, researchers can determine if gaps exist in their validated SNP collection for the region of interest. SNPbrowser's SNP density wizard assists in filling in these gaps by facilitating SNP selections that are evenly spaced between genetic locations and density within regions (Figure 4). This "gap filling" functionality dramatically reduces the time spent manually selecting SNPs to meet density requirements. SNP selection is based upon spacing of markers and can be performed with either LD or physical distance (LDUs or kilobases). Selecting evenly spaced SNPs using LDUs identifies a set

of markers with equal coverage of all regions based on the LD distribution. Thus, fewer markers are selected in regions where LD is high and more markers are selected where historical recombination events have eroded useful LD for mapping. Alternatively, you can select SNPs in a picket-fence pattern using kilobase spacing (Figure 4).

Power Calculations

To further simplify study design and to find a disease susceptibility mutation, statistical power information is included with the software. Statistical power is used for finding disease susceptibility mutations, using SNP genotyping data generated by either Applied Biosystems or HapMap databases. It has been estimated based on the empirical genotype data obtained from the respective reference population panels. Statistical power is provided on a per gene basis, including 10 kb up and downstream of the transcribed region. Power is represented by a color scale defined in the top right-hand corner of the screen, and is shown by the different population bars. Power calculations are extremely useful at the beginning of association studies to answer questions like:

- Do you have enough SNPs for the genes that you're trying to study?
- Do you need more samples?
- Where do you need to supplement with additional SNPs?

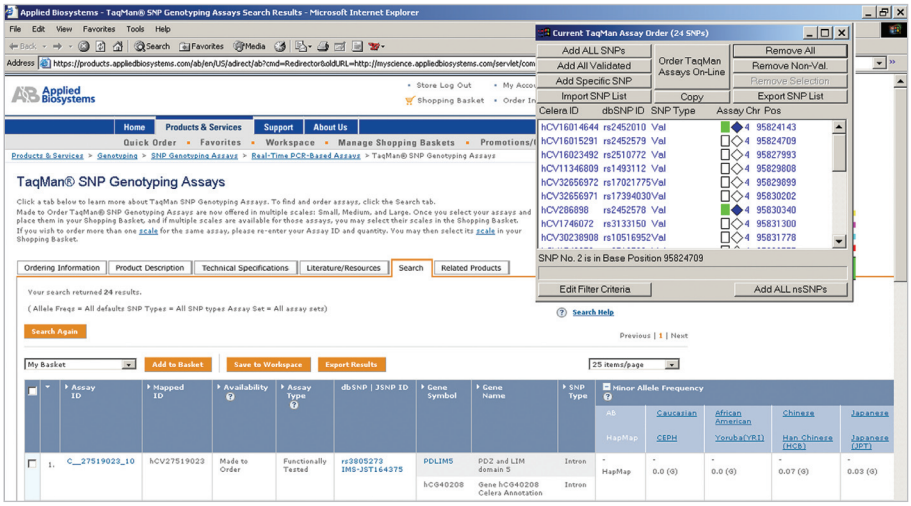

Figure 5. Connect directly to Applied Biosystems to purchase assays associated with SNPs selected using SNPbrowser Software.

Simple Online Ordering

SNPbrowser Software v3.5 contains over six million SNPs including product associations with TaqMan SNP Genotyping Assays and the SNPlex Genotyping System. SNPbrowser now includes all high quality International HapMap Project SNPs and an extensive offering of non-synonymous TaqMan® Coding SNP (cSNPs) Assays. These cSNP assays are designed to identify genetic variations that cause changes at the protein level. These variations are more likely than other variations in the non-coding regions to cause predisposition to disease or influence positive or negative response to drug therapy. The collection represents the most complete pre-designed, genome-wide assay set based on the recently completed International HapMap data set. This project has developed a public resource to help researchers find genes associated with human disease and response to pharmaceuticals. SNPbrowser supports a direct link to the Applied Biosystems online store to easily facilitate ordering of desired SNP assays (Figure 5).

For more information, and to get free SNPbrowser Software v3.5, visit www.allsnps.com/snpbrowser

Printed in the USA, 10/2006 Publication 127PB10-03

Headquarters
850 Lincoln Centre Drive | Foster City, CA 94404 USA
Phone 650.638.5800 | Toll Free 800.345.5224
www.appliedbiosystems.com

International Sales
For our office locations please call the division headquarters or refer to our Web site at
www.appliedbiosystems.com/about/offices.cfm