The
SNP Consortium Ltd.
8770 West Bryn Mawr Avenue, Chicago, IL 60631, USA • Fax: +1-773-867-2991 • Telephone: +1-773-867-2990

EMBARGOED UNTIL APRIL 15

Contact: Mary Prescott [redacted]
Jeanne Corrigan [redacted]

CONSORTIUM OF PHARMACEUTICAL COMPANIES AND THE WELLCOME TRUST FUND CREATION OF PUBLIC DATABASE OF GENE MARKERS

- Leading Academic Centers Collaborate in "SNP" Mapping -

- Improved Understanding of Biological Basis of Disease Expected To Lead to More and Better Treatments -

London and New York (April 15) – A powerful new tool to enhance understanding of disease processes and facilitate discovery, development and delivery of safer and more effective medications is the promise of an innovative collaboration among the pharmaceutical industry, the world's largest medical research charity, and several leading academic centers.

Announced today, the two-year, $45 million initiative to create a high-quality map of genetic markers is being funded by the Wellcome Trust and ten pharmaceutical companies: AstraZeneca PLC, Bayer AG, Bristol-Myers Squibb Company, F. Hoffmann-La Roche, Glaxo Wellcome PLC, Hoechst Marion Roussel AG, Novartis, Pfizer Inc, Searle, and SmithKline Beecham PLC.

Leading academic centers, including the Whitehead Institute for Biomedical Research, Washington University School of Medicine in St. Louis, the Wellcome Trust's Sanger Centre, Stanford Human Genome Center, and Cold Spring Harbor Laboratory, will participate in the identification and analysis of the genetic markers known as single nucleotide polymorphisms, or SNPs (pronounced "snips").

- more -

This collaborative effort, called The SNP Consortium, is organized as a non-profit entity. Governed by a board composed of representatives of the member organizations and led by an independent chairman, the consortium participants provide oversight and technical expertise to the scientific plan, and direct the effort to ensure public availability of SNPs that are generated.

As it is constructed, the SNP map will be placed in the public domain, giving free and equal access to all in the worldwide medical research community. The map is expected to be useful in:

- identifying specific genes involved in both common and rare diseases, thereby facilitating the discovery of new ways and new medicines to intervene in the disease process,

- developing novel diagnostic tests, and

- creating new "personalized" medicines, based on an understanding of minute genetic variations that predict response to therapy.

SNPs are common variations that occur in human DNA, and which scientists believe can help pinpoint the subtle genetic differences that predispose some but not others to diseases such as arthritis, Alzheimer's, cancer, diabetes, and depression, and underlie variability in individual responses to a given drug. In turn, such knowledge can be used to spur the development of new therapies that will better treat, or even prevent, illness.

"A large, high-density and high-quality SNP map will be of great utility to the medical research community, as it will help answer questions about genetic factors that contribute to disease susceptibility and response to treatment, and suggest directions for future investigation," says Arthur Holden, chairman and chief executive officer of The SNP Consortium. "The members of the consortium believe that free and unrestricted access to this powerful tool will benefit scientific inquiry in industry, government, academic, and independent laboratories."

Through the collaboration, it is expected that a high-density, high-quality map will be created more quickly, and with shared financial risk and less duplication of effort than if each company pursued development of a SNP map on its own. Once created, the map will provide important direction for industry scientists searching for new ways for medicines to cure and prevent disease, and for medicines that more precisely target specific patient groups.

Moreover, the availability of the map to academic, government, and independent medical researchers worldwide should enable investigation of genes associated with rare diseases, which otherwise might not be feasible because of the significant investment that would be required.

The value of SNPs resides in their simplicity, frequency, and distribution throughout the human genome (the full set of genetic instructions, encoded in long strands of DNA). There are roughly 3 billion nucleotide pairs (the "building blocks" of DNA, also called "base pairs") in the human genome, 99.9 percent of which are the same for everyone. It is the remaining 0.1 percent that accounts for differences between people. A variation in a single base – that is, a SNP – is estimated to occur every 1000 bases.

The SNP Consortium intends to identify up to 300,000 and map at least 150,000 SNPs that can then be used in association studies. That is, SNP patterns from a target population – such as patients who suffer from a particular disease or who respond poorly to a particular drug – can be compared with SNP patterns from unaffected populations to find genetic variations shared only by the affected group. It is from these association studies that disease-specific genes may be identified, and from which novel therapeutic avenues and even tailor-made therapies are expected to evolve.

- more -

Using DNA from a diversified, representative panel of anonymous volunteers, sequence information from the publicly-funded Human Genome Project, and advanced sequencing and mapping technologies, the Whitehead Institute, Washington University School of Medicine in St. Louis, the Wellcome Trust's Sanger Centre, and Stanford Human Genome Center will identify and map SNPs. The bioinformatics program at Cold Spring Harbor Laboratory will use computerized methods to organize, analyze, and manage the resulting SNP database, as well as distribute the information contained in the database. As SNPs are identified, steps will be taken to establish the date of discovery, and to guarantee free access to the map that is developed.

"The mission of The SNP Consortium is to make widely available an important research tool that will advance our understanding of disease processes, and by extension, the field of human medicine," Mr. Holden says. "We believe our efforts will be helpful to medical researchers worldwide, and will complement other public undertakings such as those spearheaded by the National Human Genome Research Institute, the Wellcome Trust, and the Medical Research Council's Human Genome Mapping Project, as well as the many private initiatives currently underway. The knowledge gained in the next several years has the potential to change fundamentally the practice of medicine."

###

The
SNP Consortium Ltd.
8770 West Bryn Mawr Avenue, Chicago, IL 60631, USA • Fax: +1-773-867-2991 • Telephone: +1-773-867-2990

SINGLE NUCLEOTIDE POLYMORPHISMS (SNPs) AND THE PERSONALIZATION OF MEDICINE

– Background Information -

Single nucleotide polymorphisms (SNPs) are likely to become important tools for genetic research. The following describes what SNPs are, how they currently are being used in human genome studies, how they likely are to be used in future research, and why it is important to make a SNP database freely accessible.

Polymorphisms

Each cell in the human body contains a full set of instructions for making a human being. These instructions, known collectively as the human genome, are encoded in long strands of DNA which are divided into genes and packed into chromosomes in the command center of each cell, known as the nucleus. Among individuals, the vast majority of DNA – more than 99 percent, in fact – is identical. Most of the characteristics that distinguish individuals from one another arise from a relatively small number of stable, frequently occurring genetic variations known as polymorphisms. Some polymorphisms likely produce innocuous differences in appearance: eye- and hair-color, for example. Other polymorphisms have been identified that affect drug metabolism, or render an individual susceptible to diseases such as cancer and Alzheimer's disease. It is thought that disease-susceptibility polymorphisms interact with environmental factors, other genes, or a combination of both to produce illness. Other polymorphisms may have no effect at all.

- more -

The most common type of genetic variation is the SNP, a change at a single DNA building-block (nucleotide). Some SNPs may contribute to disease states, but most probably do not. Still, SNPs have become essential tools for researchers attempting to decipher and create a detailed "map" of the entire human genome.

Why map the genome?

The aim of the ongoing worldwide effort to decode (sequence) and map the human genome is to provide investigators in the next century with an incredibly powerful research tool – one that will allow them to chart human evolution, study genes that interact in complex biological processes, and pinpoint genetic changes that are involved in complex illnesses such as heart disease and diabetes. The practical applications of such studies are many. For example, the identification of disease-susceptibility genes will create a host of new ways in which medicines can cure or prevent disease. This will give rise to rationally designed drug or gene therapy. Instead of merely managing symptoms, doctors will be able to use such therapeutics to prevent or cure previously intractable illnesses. Studies of genetic variation also can be used (and are being used) to identify those patients who are most likely to benefit from a given therapy, and those who are most likely to suffer debilitating or life-threatening side effects. These applications are likely to become widespread in the near future, and will allow pharmaceutical and biotechnology companies to develop – and doctors to administer – personalized therapies rather than hit-or-miss general treatments.

The role of polymorphisms (genetic markers) in genome studies

Creating a genome map involves breaking chromosomes into pieces, sequencing the fragmented DNA, and reassembling the decoded fragments into their original order. Because the human genome comprises some 3 billion nucleotide pairs (the "building blocks" of DNA, also called "base pairs") and around 100,000 genes, the mapping effort has been extremely labor-intensive.

To facilitate the process of piecing together hundreds of thousands of DNA fragments, researchers make use of genetic landmarks called "markers." Put simply, markers are easy-to-track polymorphisms – be they alternative versions of genes (alleles), repetitive stretches of DNA, or alterations in single nucleotides (SNPs). Using genetic analysis and other "physical" means, genome researchers can determine the approximate distance between two known markers, and thereby assign them to positions on a particular chromosome.

Genetic markers also help scientists to navigate the genome map; much like coordinates on a conventional map, they allow researchers to pinpoint an area of interest, and to follow that area in genetic association studies. For example, if a marker (or group of markers) is present in patients with a particular illness, but not in healthy persons, then it is very likely that the region of DNA containing the marker contributes to the illness in question. Once such an association has been established, the marker can be used to screen patients for disease-susceptibility.

Why develop SNP maps?

The current map of the human genome is dotted with many landmarks, but most of them are too far apart for finely detailed analysis: It is difficult to pinpoint a disease-related region of DNA when the nearest signposts are a million or more nucleotides away. To increase the level of detail, and therefore improve the ability to zero in on a potentially interesting gene, genome researchers have turned to SNPs. The strength of SNPs as genetic landmarks lies in their stability, their distribution, their frequency, and their simplicity:

- ***Stability.*** Unlike some markers used in earlier efforts, SNPs do not appear to change much from generation to generation, and therefore may be more reliable and easier to track.
- ***Distribution.*** Whereas conventional genetic markers appear at irregular intervals, often clustered in certain regions of DNA, SNPs are relatively evenly spaced throughout the genome.

- ***Frequency.*** Unlike markers on existing maps, which are often separated by a million or more nucleotides, SNPs appear throughout the genome at an average of once every 1000 nucleotides.
- ***Simplicity.*** SNPs are the simplest form of polymorphism, and are therefore easy to detect. In fact, they are ideal for use in currently available, digitized DNA-analyzing systems, which permit rapid, automated screening of thousands of SNPs at once.

The assembly of dense arrays of high-quality SNPs across the existing genome map likely will provide future investigators with a research tool of unprecedented power – one that may help to identify the molecular basis of devastating diseases, both common and rare; one that may lead to new therapies that not only control symptoms, but also cure the underlying illness; one that may assist in accurate diagnosis and early or preventive treatment; and one that may help to match the right treatment to the right patient. Collectively, such advances are expected to bring about an era of personalized medicine.

Public versus private SNPs

Given the potential importance of SNPs as a research tool, it is not surprising that there is considerable interest, both public and private, in the development of high-quality SNP maps. However, because SNP identification is but a very early step in the research process, efforts to map new SNPs are truly pre-competitive. Indeed, the pharmaceutical companies in The SNP Consortium believe that making dense, high-quality maps freely and widely available benefits everyone – commercial, academic, and government research centers; and, especially, the patients who will profit from better therapies. This view is shared by the National Human Genome Research Institute (NHGRI); as such, both the NHGRI and The SNP Consortium – a partnership among charitable, for-profit, and academic organizations – have pledged to create publicly available and freely accessible repositories of SNPs. Such repositories will increase in value as new discoveries are "annotated" to the existing map.

In summary, SNPs are essential tools both for present genome-mapping efforts and for important studies to come. The sooner dense, high-quality SNP maps become available to the research community at large – which includes pharmaceutical and biotechnology companies, academic and clinical laboratories, and government health organizations – the sooner innovative investigations can proceed in all quarters. The results of such investigations could radically change the way medicine is practiced, shifting the focus of treatment from what works best for the majority to what works best for each individual.

###

GLOSSARY

Allele:

One of two or more alternative versions of a gene; for example, different versions of the gene for β-hemoglobin are either normal (called "wild-type") or sickle-cell (causing disease).

Association Genetics:

A means of establishing associations between specific genes and specific diseases, requiring a large clinical database containing both medical and DNA information for each patient and a control group.

Base Pair:

Two nitrogenous bases (adenine and thymine or guanine and cytosine) held together by weak bonds. Two strands of DNA are held together in the shape of a double helix by the bonds between base pairs. See **nucleotide.**

Bioinformatics:

The science that uses advanced computing techniques for management and analysis of biological data. Bioinformatics is particularly important as an adjunct to genomic research, which generates a large amount of complex data, involving billions of individual DNA building-blocks, and tens of thousands of genes.

Chromosome:

Structures found in the nucleus of a cell, which contain genes. Chromosomes come in pairs, and a normal human cell contains 46 chromosomes, 22 pairs of autosomes and two sex chromosomes.

Coding region:

The part of a gene that specifies the structure of a protein.

Cloning:

The generation of a cell or organism that is genetically identical to another cell or organism. In molecular biology, cloning refers to the act of isolating a piece of DNA from an organism and inserting it into another DNA molecule (vector) that is capable of replicating in bacteria or other cell-types. The vector containing the inserted piece of DNA is referred to as a **clone.**

DNA (deoxyribonucleic acid):
The molecule that carries the genetic information (blueprint) for most living systems. DNA is a linear, double-stranded chain of nucleotides that is packed into chromosomes in the nucleus of every cell.

DNA bank:
A storage repository for DNA extracted from blood samples or other human tissue. Analysis of DNA samples stored in the DNA bank may be used to help individuals and their physicians trace the pattern of disease in families, or to aid future medical research areas.

DNA library:
A collection of cloned DNA fragments from an organism's genome. A particular library may include clones of all the DNA sequences expressed in a certain kind of cell, or in a certain organ of the body, or fragments from the nuclear DNA (i.e., the genome) of an organism.

Double helix:
The shape that two individual strands of DNA assume when bonded together by base pairing. See **base pair.**

Gene:
The basic functional unit of heredity located on a chromosome. Genes are blueprints for proteins, which are central to all life-processes.

Gene mapping:
Determination of the relative location of a gene on a chromosome.

Genetics:
The study of heredity: how biological information is transferred from one generation to the next as well as how that information is expressed within an organism.

Genetic predisposition:
Susceptibility of an organism to develop a specific disease due to genetic alterations (mutations). Because environmental factors play a large part in the expression of certain traits, the organism may or may not actually develop the predicted disease.

Genetic testing:
Analyzing the hereditary profile of an organism; the use of specific biologic assays to analyze an individual's genetic profile.

- more -

Genome:
An organism's complete set of chromosomes and genes; the total hereditary material in an organism.

Genome map:
A reconstruction of the entire set of chromosomes for a given organism, showing the relative position of every gene.

Genomics:
The study of an organism's entire complement of genetic material.

Genotype:
The genetic characteristics or description of an organism defined by the nucleotide sequence of the genome.

Human Genome Project:
An ongoing, global research effort aimed at mapping and decoding the complete human genome.

Marker:
A commonly-occurring genetic variation that can be easily tracked in genetic studies. Markers can be entire alleles, repetitive stretches of DNA, or single nucleotide polymorphisms.

Mutation:
An alteration in a gene that can be transmitted from one generation to the next. Although many mutations are associated with defects, some have no effect on health of an organism.

Non-coding region:
The part of a gene that does not specify the structure of a protein. Non-coding regions of DNA often contain elements that regulate when a protein will be made, and how much of that protein will be produced.

Nucleotide:
The basic building block of DNA, composed of a sugar, a phosphate group, and a nitrogenous base (adenine, guanine, cytosine, or thymine). See **base pair**.

Pharmacogenomics:
The application of genomics to pharmaceutical research; using genome studies to identify genes that affect how drugs work in different individuals.

- more -

Phenotype:
The physical appearance or observed characteristics of an organism that are the result of the interaction between the genetic characteristics (genotype) of the organism and its environment.

Polymorphism:
A commonly occurring variation in a gene or segment of DNA; also, the existence of several forms of a gene (alleles) in the same species.

Protein:
A large molecule composed of one or more chains of amino acids in a specific order; the order is determined by the base sequence of nucleotides in the gene coding for the protein. Proteins are required for the structure, function, and regulation of the body's cells, tissues, and organs, and each protein has unique functions. Examples are hormones, enzymes, and antibodies.

RNA (ribonucleic acid):
Chemical found in the nucleus and cytoplasm of cells; it plays an important role in protein synthesis and other chemical activities of the cell. The structure of RNA is similar to that of DNA, although RNA is single stranded whereas DNA has a characteristic "double helix."

Sequencing:
The process of determining the specific order of nucleotides in a DNA molecule. Sequencing also refers to determining the order of amino acids in a protein.

Single nucleotide polymorphism (SNP):
The most common type of genetic variation, consisting of a change at a single base in a DNA molecule.

SNP map:
A collection of SNPs that can be superimposed over the existing genome map, creating greater detail, and facilitating further genetic studies.

###