Review

Bench-to-bedside review: Fulfilling promises of the Human Genome Project

Jean-Daniel Chiche[1], Alain Cariou[2] and Jean-Paul Mira[3]

[1]Associate Professor of Critical Care Medicine, Medical Intensive Care Unit and Cochin Institute of Molecular Genetics, Hôpital Cochin, Université René Descartes, Paris, France
[2]Associate Professor of Critical Care Medicine, Medical Intensive Care Unit, Hôpital Cochin, Université René Descartes, Paris, France
[3]Professor of Critical Care Medicine, Medical Intensive Care Unit and Cochin Institute of Molecular Genetics, Hôpital Cochin, Université René Descartes, Paris, France

Correspondence: Jean-Daniel Chiche, jean-daniel.chiche@cch.ap-hop-paris.fr

Published online: 20 March 2002 *Critical Care* 2002, **6**:212-215
© 2002 BioMed Central Ltd (Print ISSN 1364-8535; Online ISSN 1466-609X)

Abstract

Since most common diseases have been shown to be influenced by inherited variations in our genes, completion of the Human Genome Project and mapping of the human genome single-nucleotide polymorphisms will have a tremendous impact on our approach to medicine. New developments in genotyping techniques and bioinformatics, enabling detection of single-nucleotide polymorphisms, already provide physicians and scientists with tools that change our understanding of human biology. In the near future, studies will relate genetic polymorphisms to features of critical illnesses, increased susceptibility to common diseases, and altered response to therapy. Novel insights into the contribution of genetic factors to critical illnesses and advances in pharmacogenomics will be used to select the most effective therapeutic agent and the optimal dosage required to elicit the expected drug response for a given individual. Implementation of genetic criteria for patient selection and individual assessment of the risks and benefits of treatment emerges as a major challenge to the pharmaceutical industry.

Keywords genetics, pharmacogenomics, polymorphism

In 1995, the genomic sequence of the bacteria *Haemophilus influenzae* was the first complete genomic sequence of a free-living organism to be published [1]. Since then, scientists have totally sequenced the genomes of more than one hundred bacteria and completed genetic maps of large multicellular organisms [2–5]. The draft sequence of the human genome, recently published by the Human Genome Project public consortium [6] and by a private company [7], represents a milestone in science. Today, the genetic blueprint for a human is nearly completed and covers 96% of the genome. Embedded within our genomes are the sequences of the approximately 30,000 genes that underlie human biology and medicine. As we enter the post genome-sequencing era, we are already facing new challenges. Successful translation of this structural knowledge into clinical benefits will depend upon our ability to relate individual genes to specific diseases, to find the genetic variations that influence an individual's risk of becoming ill, and to use genetic information to tailor drug therapy. The purpose of this review is to put some of the predictable consequences of the advances in genomics into clinical perspective.

Single-nucleotide polymorphisms: learning from our differences

Most common diseases and many drug responses have been shown to be influenced by inherited differences in our genes. Thus, studying generic variance can improve our understanding and treatment of disease. If a region of the human genome is sequenced from two randomly chosen individuals, 99.3% of the examined DNA will be identical [8]. Much of the genetic variation between individuals lies in differences known as single-nucleotide polymorphisms (SNPs); a single base is swapped for an alternate, and both versions exist in the general population at frequencies greater than 1% [8]. As SNPs constitute the bulk of human genetic variation, they can be used to track inheritance of genes in traditional family-

SNP = single-nucleotide polymorphism, TNF = tumor necrosis factor.

based linkage studies. By epidemiological association, SNPs can also be used to test susceptibilities to common diseases such as heart disease, cancer, and diabetes.

Based on the promise of SNP research, an international subset of academic centers, pharmaceutical companies, and a private foundation teamed up to create the SNP Consortium in 1999. Whereas the initial goal of the consortium was to discover 300,000 SNPs that would be freely available by April 2001, this has been exceeded, and the SNP, in collaboration with the International Human Genome Sequencing Consortium, has created a catalogue of more than 1.4 million SNPs [9]. This publicly available SNP map promises to advance our knowledge of the links between genes and diseases.

Linking SNPs to phenotypes: disease markers or more?

One of the most difficult challenges faced by physicians and scientists is to establish the link between gene variations and a disease. Of the 1.4 million SNPs currently on the public map, only 60,000 are located in protein coding regions, called exons, and relatively few of these transform amino acids [9]. The SNPs that change the amino acid sequence, and variants in gene regulatory regions that control protein expression levels, are most likely to have a direct impact on the protein product of a gene [10]. In cases where change of a single base in the genome sequence is sufficient to cause disease, it has become possible to identify this change and improve our understanding of the disease. For instance, sickle cell anemia is caused by the substitution of a thymine for adenine at a single position in the gene that encodes the hemoglobin molecule.

Using ever more powerful approaches, literally hundreds of rare human diseases have been related to genetic defects. However, the genetic contributions have proven more difficult to establish for the common diseases that account for most morbidity and mortality. In most cases, the influence of gene variants is subtle and the risk of contracting the disease is also influenced by environmental factors [10]. Even if the causal mutations are common in the population, their effects will, therefore, be difficult to discover. As the effects of any given SNP may be modest, it will be necessary to study large numbers of patient samples to observe associations in a reproducible fashion. Therefore, comprehensive studies will rely on the development of fast and efficient tools to identify the small number of relevant SNPs out of the millions in the human genome.

A phenomenon called linkage disequilibrium should permit the use of SNPs to track associations to disease, without necessarily finding each functionally relevant SNP beforehand. In a certain region, SNPs often track together in the population. In linkage disequilibrium, such nearby SNPs can serve as proxies for each other in a disease study. Hence, a subset of SNPs spaced throughout the genome might allow a comprehensive test of common genetic variation across the entire genome. Although the specific number of SNPs needed for linkage disequilibrium studies is unknown, the 1.4 million SNPs in the public domain should offer a sufficient number to explore most regions of the genome.

Impact of SNP research on clinical trial design

Besides all the consequences of genetics on our understanding of the pathophysiology of critical illnesses, advances in SNP research also promise to change current practices in clinical trials [11]. The SNP effort will undoubtedly serve as the bedrock of pharmacogenomics, the emerging field of personalized medicine in which drugs and preventative strategies are specifically tailored to suit an individual's genetic profile. One can speculate that many of the recent advances in genetics will soon be brought into clinical trials with two main directions. First, whereas treatment allocation has been based mainly on phenotype, genetic characterization based on the genetic profile of an individual will help researchers to identify suitable subjects to test a working hypothesis. This approach will also facilitate interpretation of the results of clinical trials, and ultimately enable clinicians to tailor treatment to patients with specific genotype. For instance, an analysis based on the main studies of anti-tumor necrosis factor (TNF) strategies in septic patients found an absolute decrease in mortality of 3.5%, suggesting that these therapies could be beneficial in septic patients with uncontrolled TNF release [12]. Targeting patients whom carry the TNF2 allele and produce high levels of TNF-α [13], may reveal a beneficial effect of treatment with anti-TNF antibodies for septic shock [14,15]. Second, as interindividual variability in the response to drugs remains a substantial clinical problem, a major objective of pharmacogenomic research is to decrease adverse responses to therapy through determination of adequate therapeutic targets and genetic polymorphisms that alter drug specificity, metabolism, and toxicity [11]. Ultimately, genetic information will be used to select the most effective therapeutic agent and the optimal regimen to elicit the expected drug response for a given individual. Hence, the implementation of genetic criteria to select patient populations and of individual assessment of the risks and benefits of treatment is emerging as a major challenge for pharmaceutical companies.

Among the hurdles to overcome for successful integration of genetics in clinical practices, it will be necessary to improve our ability to detect SNPs at a lower cost. Methods to identify SNPs are based on modifications of the traditional DNA sequencing approach, which can use a range of detection methods, such as radioactivity, fluorescence resonance energy transfer, or fluorescence polarization. More recently, arrays on glass slides, DNA chip-based microarrays, and mass spectrometry genotyping technologies have been introduced to simultaneously determine the genotype of large numbers of SNPs [16–18]. It is not yet clear which of these

powerful methods will become most useful. At a current average price of one dollar per genotype, SNP detection in large-scale genotyping studies is still prohibitively expensive. Even at one cent per genotype, the cost per patient in a typical association study testing 100,000 SNPs will possibly add one million US dollars to the cost of a clinical trial [19]. Significant advances will be necessary to make extensive genotyping a standard part of clinical trials.

Perspectives and limitations of SNP research

There are still many significant technical and analytical problems that must be solved before the promise of SNPs can be fulfilled. Whereas the current SNP maps provide us with invaluable tools to track statistically significant associations between SNPs and disease or drug response, we do not fully understand the genetic architecture of common traits underlying disease susceptibility and variability in drug response. Interpretation of association studies is complicated by the number of genes, the number of variants in each gene, and the frequency of a variant within a population. Location of a variant SNP in the coding region, the regulatory region, or the noncoding region of the genome also affects susceptibility to disease in a way that is still unclear. In addition, the interaction of individual SNPs and the degree to which they track together in linkage disequilibrium may be of the utmost importance in the determination of a given phenotype.

Other issues must be addressed to unlock the full potential of SNPs. Given the large number of SNPs and the low probability that any specific one causes disease, the sample sizes in association studies need to be large enough to achieve adequate statistical power. This also raises the problem of accurately phenotyping individuals, since the same disease may manifest itself with different patterns in different patients. Finally, new ethical issues will arise, which will have to be solved as SNP technology improves and becomes widely used. Whereas current genetic tests typically track single-disease genes, SNPs will provide tests that associate a genetic profile with individual predisposition to a broad list of diseases. Physicians and scientists are just beginning to address the question of how to keep such sensitive phenotypical and genotypical information confidential so that it is not misused by either employers or insurance companies. Most importantly, our patients will have to cope with this information, sometimes left in the expectation of preventive strategies and therapeutic solutions.

As the first round of human genome sequencing nears completion, identifying functions for each of the 30,000 or more human genes, and determining which of these genes play a role in disease, will emerge as one of the great challenges of twenty-first century biomedicine. Yet, physicians and scientists have undertaken the task of characterizing and cataloging a shared universe of generic differences that underlies our susceptibility to diseases and alters our response to drugs. Although this work appears to be quite demanding, it provides tremendous opportunities in our search to understand, and ultimately treat, diseases that account for most of the mortality and morbidity in our intensive care units.

Competing interests

None declared.

References

1. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM, *et al.*: **Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.** *Science* 1995, **269**:496-512.
2. McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L, Porwollik S, Ali J, Dante M, Du F, Hou S, Layman D, Leonard S, Nguyen C, Scott K, Holmes A, Grewal N, Mulvaney E, Ryan E, Sun H, Florea L, Miller W, Stoneking T, Nhan M, Waterston R, Wilson RK: **Complete genome sequence of Salmonella enterica serovar Typhimurium LT2.** *Nature* 2001, **413**:852-856.
3. Parkhill J, Wren BW, Thomson NR, Titball RW, Holden MT, Prentice MB, Sebaihia M, James KD, Churcher C, Mungall KL, Baker S, Basham D, Bentley SD, Brooks K, Cerdeno-Tarraga AM, Chillingworth T, Cronin A, Davies RM, Davis P, Dougan G, Feltwell T, Hamlin N, Holroyd S, Jagels K, Karlyshev AV, Leather S, Moule S, Oyston PC, Quail M, Rutherford K, Simmonds M, Skelton J, Stevens K, Whitehead S, Barrell BG: **Genome sequence of Yersinia pestis, the causative agent of plague.** *Nature* 2001, **413**:523-527.
4. Parkhill J, Dougan G, James KD, Thomson NR, Pickard D, Wain J, Churcher C, Mungall KL, Bentley SD, Holden MT, Sebaihia M, Baker S, Basham D, Brooks K, Chillingworth T, Connerton P, Cronin A, Davis P, Davies RM, Dowd L, White N, Farrar J, Feltwell T, Hamlin N, Haque A, Hien TT, Holroyd S, Jagels K, Krogh A, Larsen TS, Leather S, Moule S, O'Gaora P, Parry C, Quail M, Rutherford K, Simmonds M, Skelton J, Stevens K, Whitehead S, Barrell BG: **Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18.** *Nature* 2001, **413**:848-852.
5. Gaasterland T, Oprea M: **Whole-genome analysis: annotations and updates.** *Curr Opin Struct Biol* 2001, **11**:377-381.
6. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blocker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T,

Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kaspryzk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, Szustakowki J; International Human Genome Sequencing Consortium: **Initial sequencing and analysis of the human genome.** *Nature* 2001, **409**:860-921.

7. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferriera S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratts E, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers YH, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J, Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigo R, Campbell MJ, Sjolander KV, Karlak B, Kejariwal A, Mi H, Lazareva B, Hatton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooseph S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang YH, Coyne M, Dahlke C, Mays A, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M, Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A, Zhu X: **The sequence of the human genome.** *Science* 2001, **291**:1304-1351.
8. Collins FS, Guyer MS, Charkravarti A: **Variations on a theme: cataloging human DNA sequence variation.** *Science* 1997, **278**:1580-1581.
9. Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, Sherry S, Mullikin JC, Mortimore BJ, Willey DL, Hunt SE, Cole CG, Coggill PC, Rice CM, Ning Z, Rogers J, Bentley DR, Kwok PY, Mardis ER, Yeh RT, Schultz B, Cook L, Davenport R, Dante M, Fulton L, Hillier L, Waterston RH, McPherson JD, Gilman B, Schaffner S, Van Etten WJ, Reich D, Higgins J, Daly MJ, Blumenstiel B, Baldwin J, Stange-Thomann N, Zody MC, Linton L, Lander ES, Altshuler D; The International SNP Map Working Group: **A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms.** *Nature* 2001, **409**:928-933.
10. Subramanian G, Adams MD, Venter JC, Broder S: **Implications of the human genome for understanding human biology and medicine.** *JAMA* 2001, **286**:2296-2307.
11. Phillips KA, Veenstra DL, Oren E, Lee JK, Sadee W: **Potential role of pharmacogenomics in reducing adverse drug reactions: a systematic review.** *JAMA* 2001, **286**:2270-2279.
12. Marshall JC: **Clinical trials of mediator-directed therapy in sepsis: what have we learned?** *Intensive Care Med* 2000, **26**: S75-S83.
13. Mira JP, Cariou A, Grall F, Delclaux C, Losser MR, Heshmati F, Cheval C, Monchi M, Teboul JL, Riche F, Leleu G, Arbibe L, Mignon A, Delpech M, Dhainaut JF: **Association of TNF2, a TNF-alpha promoter polymorphism, with septic shock susceptibility and mortality: a multicenter study.** *JAMA* 1999, **282**: 561-568.
14. Fisher CJ Jr, Opal SM, Dhainaut JF, Stephens S, Zimmerman JL, Nightingale P, Harris SJ, Schein RM, Panacek EA, Vincent JL, *et al.*: **Influence of an anti-tumor necrosis factor monoclonal antibody on cytokine levels in patients with sepsis. The CB0006 Sepsis Syndrome Study Group.** *Crit Care Med* 1993, **21**:318-327.
15. Abraham E, Wunderink R, Silverman H, Perl TM, Nasraway S, Levy H, Bone R, Wenzel RP, Balk R, Allred R, *et al.*: **Efficacy and safety of monoclonal antibody to human tumor necrosis factor alpha in patients with sepsis syndrome. A randomized, controlled, double-blind, multicenter clinical trial. TNF-alpha MAb Sepsis Study Group.** *JAMA* 1995, **273**:934-941.
16. Hirschhorn JN, Sklar P, Lindblad-Toh K, Lim YM, Ruiz-Gutierrez M, Bolk S, Langhorst B, Schaffner S, Winchester E, Lander ES: **SBE-TAGS: an array-based method for efficient single-nucleotide polymorphism genotyping.** *Proc Natl Acad Sci USA* 2000, **97**:12164-12169.
17. Fan JB, Chen X, Halushka MK, Berno A, Huang X, Ryder T, Lipshutz RJ, Lockhart DJ, Chakravarti A: **Parallel genotyping of human SNPs using generic high-density oligonucleotide tag arrays.** *Genome Res* 2000, **10**:853-860.
18. Bray MS, Boerwinkle E, Doris PA: **High-throughput multiplex SNP genotyping with MALDI-TOF mass spectrometry: practice, problems and promise.** *Hum Mutat* 2001, **17**:296-304.
19. McCarthy JJ, Hilfiker R: **The use of single-nucleotide polymorphism maps in pharmacogenomics.** *Nat Biotechnol* 2000, **18**: 505-508.