Allergology International (2004) **53**: 87–92

Review Article

Translation of the human genome into clinical allergy, part 2

Hirohisa Saito

Department of Allergy and Immunology, National Research Institute for Child Health and Development and Laboratory for Allergy Transcriptome at RIKEN Research Center for Allergy and Immunology, Tokyo, Japan

Abstract

After completion of sequencing of the human genome, you will no longer be able to discover a new gene and may not be able to find a new molecule in the human body. Instead, you may find many new molecule networks. It will be possible to select information obtained from animal models just where orthologous genes are functioning similarly. Mouse disease models will not be used any longer where key orthologous genes are working differently than in humans. Analysis of cell type-selective transcripts from database searches is now available to minimize the efforts required for drug discovery. As such, it will soon be possible to use computational modeling to analyze integrative biological function and to test hypotheses without performing any *in vivo* or *in vitro* experimentation. However, before establishing a system simulating the human body, which consists of a variety of organs, which further consist of various types of cells, which then consist of various types of proteins, which consist of 20 types of amino acids, there are many steps that need to be understood.

Key words: basophils, mast cells, microarray, system biology, transcriptome.

Correspondence: Dr Hirohisa Saito, Director, Department of Allergy and Immunology, National Research Institute for Child Health & Development and Laboratory for Allergy Transcriptome at RIKEN Research Center for Allergy & Immunology, 3-35-31 Taishido, Setagaya-ku, Tokyo 154-8567, Japan.
Email: hsaito@nch.go.jp

Received 5 January 2004.

The dawn of system biology upon clinical allergy

Being able to completely read the human genome sequence, we will know all about the role of genomic DNA sequence variations among individuals, such as a single nucleotide polymorphism, in the pathogenesis of diseases and responses to drugs. An understanding of genome will also accelerate an understanding of the transcriptome and the proteome, the functional elements in a cell. We will not miss any messages induced by a new therapy, such as an unexpected adverse effect, using such a comprehensive assay.[1] As such, the sequencing of the human genome is offering unprecedented 'system biology', aiming for the understanding of the total function of a cell, an organ or, in the near future, the human body. An understanding of the genome may provide a framework for modeling the human body in the near future, using computational methods. It may be possible to use computational modeling to analyze integrative biological function and test hypotheses.[2]

There are several different types of approaches to predict total cellular function as a consequence of the interaction of all molecules present in a cell. However, most of the approaches are simulating the final cellular function based on descriptions available in the literature and not on whole genomic information. This is because most genomic techniques, such as proteomics, are still in their infancy and are impractical, except for transcriptome analysis using microarray technology.[3–5] We have recently published two papers, which were performed by mainly analyzing our transcriptome database based on a high-density oligonucleotide microarray (GeneChip; Affymetrix, Santa Clara, CA, USA). One paper dealt with interspecies comparisons between human and mouse mast cell transcriptomes,[4] whereas the other paper considered the pharmaceutical development of anti-allergic

drugs.[5] Herein, I introduce a few examples of the studies dealing with genome-wide information, mainly by expanding the discussion of the two previously published papers.[4,5]

Interspecies comparisons

Animal disease models have been used as surrogates for humans and have been informative. The use of mouse models for diseases related to allergy and immunology has increased markedly because of the rapidly developing technologies and immunological tools to block specific pathways or to selectively knock out genes that are important for processes that contribute to the pathogenesis of the disease. By using these technologies, the genes responsible for several types of severe combined immunodeficiency and common variable immunodeficiency have been found. Controversy does exist, however, as to the relevance of these models of allergic diseases, such as in asthma.[6,7] Clinical trials sometimes fail because of the fact that the results obtained in animal studies cannot be reproduced in humans. For instance, anti-interleukin (IL)-5 (eosinophil-growth factor) antibody completely blocked the airway hypersensitivity related to eosinophil inflammation in experimental animal models of asthma.[8] However, the therapeutic application of humanized anti-IL-5 antibody did not improve the bronchial hypersensitivity of asthmatics, despite a marked decrease in eosinophil number.[9] One of the reasons why allergic disease models sometimes fail to reproduce human diseases is the complexity of human allergic diseases. Total loss of gene function usually induces severe symptoms, but not the mild symptoms seen in multifactorial diseases such as allergic diseases.[1] We have not yet succeeded in producing even the simplest mouse model to test the hygiene hypothesis to determine whether low, but not high, levels of endotoxin in an environment induce allergic diseases.

We have examined genome-wide gene expression in cultured human and mouse mast cells (triggering IgE-mediated allergic reactions) to find molecules similarly regulated and expressed by the two mast cell types.[4] Rodent mast cells are common experimental tools, but are somewhat different from their human counterparts in their responses to certain cytokines[10] and anti-allergic drugs.[11]

After stimulation via high-affinity IgE receptor (FcεRI), the transcriptional levels of several CC chemokines were markedly increased and I-309 (CCL1), macrophage inflammatory protein (MIP)-1α (CCL3) and MIP-1β (CCL4) were found among the 10 most-increased human and mouse transcripts from approximately 12 000 genes and expressed sequence tags.[4] High expression of CC chemokines by both types of activated mast cell was confirmed at the protein level using ELISA. These results suggests that mast cells play a crucial role in the recruitment of various CCR-expressing cells into the tissue in an IgE-dependent manner and that FcεRI-mediated induction of several CC chemokines is highly conserved between humans and mice.[4]

Recently, many human and mouse orthologous genes have become available at a genome-wide level in electronic format (http://www.ncbi.nlm.nih.gov/Homology/), which facilitates interspecies comparisons.[12] However, it has not been proven whether these structure-based orthologs are similarly regulated. Among these orthologous genes, we found that major basic protein mRNA levels were abundantly expressed by human mast cells, but not by mouse mast cells.[13] As such, some mRNA expression was regulated differentially in mouse and human mast cells (Table 1; http://www.nch.go.jp/imal/Mast/Blood_Dec1_2002sup.xls). Therefore, studies on the function of molecules highly expressed only in mouse cells have to be interpreted with care with regard to their potential function in humans. Interspecies comparison studies of the expression of the whole genome should be useful for the interpretation of experimental data from animal models of human pathogenesis.

To simulate the human body in a computational model, many animal models still have to be used. In the near future, it will be possible to select information obtained from animal models just where orthologous genes are functioning similarly. Mouse disease models will not be used any longer where key orthologous genes are working differently than in humans, as shown in Fig. 1.

'Druggable' genes for allergic diseases

Three types of circulating blood granulocytes (eosinophils, basophils and neutrophils) and tissue mast cells play roles in protecting against microbial infection by releasing cell type-specific mediators and proteases. Eosinophils, basophils and mast cells evoke allergic reactions, as well as damaging nematodes.[14] In addition to killing bacteria, neutrophils sometimes induce systemic vasculitis or multiple organ damage under certain conditions.[15,16] Thus, targeting granulocyte type-selective

Table 1 Representative orthologous genes that are regulated similarly or differently in human and mouse cells (http://www.nch.go.jp/imal/Mast/Blood_Dec1_2002sup.xls)

Transcripts regulated similarly in human and mouse cells
Tryptase
Heat shock 90 kDa protein
CCL2, monocyte chemotactic peptide (MCP)-1
Cathepsin D
Granzyme B
Carboxypeptidase A
CCL1, I-309
Cyclophilin B
Galectin-3
CCL4, macrophage inflammatory protein (MIP)-1β
Osteopontin
L-Histidine decarboxylase
Heat shock 60 kDa protein 1 (chaperonin)
Peroxiredoxin 1
Macrophage capping protein G-actin
Transcripts highly expressed only by human cells
Eosinophil major basic protein
Clusterin
Monoamine oxidase A
Macrophage migration inhibitory factor
Prostaglandin D_2 synthase
15-Hydroxyprostaglandin dehydrogenase
Matrix metalloproteinase 9, type IV collagenase
GM-CSF
Syndecan 3
Siglec 6
Transcripts highly expressed only by mouse cells
Chondroitin proteoglycan
GM-CSF receptor β chain
Advillin
Thrombin receptor
Neutrophil cytosolic factor 4
CD100, semaphorin
STAT5b
TGF-β
Carbonic anhydrase II
Cathepsin B

MCP, monocyte chemotactic protein; MIP, macrophage inflammatory protein; GM-CSF, granulocyte–macrophage colony stimulating factor; STAT, signal transducers and activators of transcription; TGF, transforming growth factor.

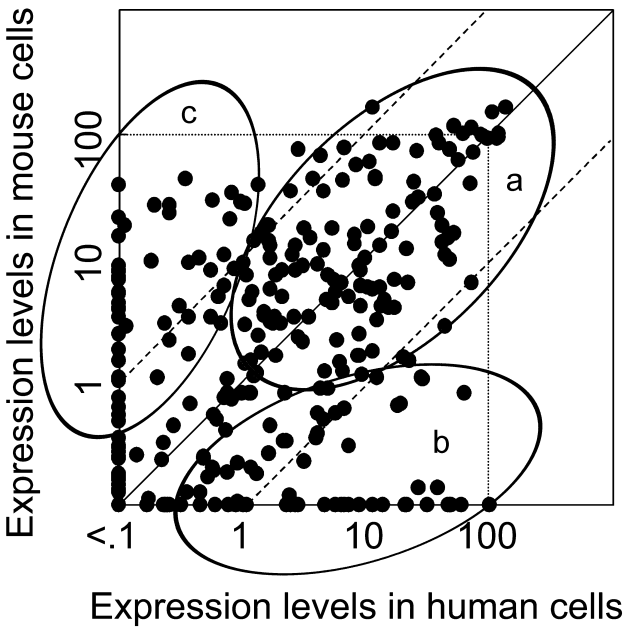

Fig. 1 Orthologous genes that are regulated similarly or differently in human and mouse cells. The expression levels of 287 pairs of orthologous genes were compared. Each point represents the average of two (human) or three (mouse) independent experiments shown at http://www.nch.go.jp/imal/Mast/Blood_Dec1_2002sup.xls. You may use mouse disease models if your research is investigating transcripts similarly regulated in human and mouse cells (a). However, if the transcripts are highly expressed only by human cells (b), you have to use human cells or other animal models. If you found that the transcripts were highly expressed only by mouse cells, you should be careful in continuing your research (c).

functions is considered an important strategy for drug discovery. Activation of these cells is generally characterized by an influx of extracellular calcium (Ca^{2+}), which is essential for the subsequent release of granule-derived mediators, newly generated lipid mediators and cytokines.[17] The flux of other ions plays an important role during granulocyte responses because these ions regulate cell membrane potential and, thus, influence Ca^{2+} influx.[18] Treatment of mast cells and basophils with pertussis toxin inactivates G_i-proteins and abolishes degranulation, but not the influx of Ca^{2+}, induced by non-immunological ligands such as thrombin and *N*-formylpeptide.[19] Thus, the granulocyte degranulation pathway is sometimes Ca^{2+} independent and is G-protein dependent. Indeed, the thrombin-activated receptors and formylpeptide receptors are classified as G-protein-coupled receptors (GPR), having a seven transmembrane region.[20] As such, ion channels and GPR both play essential roles in degranulation, as well as other cellular functions important for granulocytes, and are thought to be good targets for drug development.[21] Receptor genes and ion channel genes are found only in 5 and 1.3% of all genes present in the human genome, respectively.[22] However, receptors and ion channels are, respectively, found in 45 and 5% of the molecular targets of all known drugs.[21,23,24] In addition to their

physiological importance, receptors, including GPR and those for ion channels, are considered to be marketable and targeting these molecules should be efficient for pharmaceutical development. That is, we can concentrate on approximately 2000 genes as practical drug targets and can forget about the 30 000 other genes present in the human genome.

We used GeneChip (U133A; Affymetrix) to examine the cell type-selective transcriptome expression of seven types of leukocytes (basophils, eosinophils, neutrophils, $CD4^+$ cells, $CD8^+$ cells, $CD14^+$ cells and $CD19^+$ cells), platelets, mast cells and fibroblasts. Then, we focused on the expression of granulocyte-selective genes for ion channels, GPR and other receptors. We identified 51 granulocyte-selective genes for ion channels and receptors by examining approximately 20 000 types of transcripts derived from 16 000 genes (approximately half the number of genes present in the human genome) from the 10 different types of cells.[5] Twenty-two mast cell-, basophil- and/or eosinophil-selective transcripts identified in this study could be potential therapeutic targets for allergic diseases (Table 2)[5] because these granulocytes play a crucial role in allergic inflammation.[14] It is estimated that approximately 50 genes are selectively expressed by mast cells, eosinophils and basophils among all the genes present in the human genome.

The safety of any candidate drug must be evaluated by comparing its efficacy on these granulocytes with its toxicity to physiologically important organs. Among the 22 receptors and ion channels, three molecules were highly expressed by multiple normal tissue cell types (Table 2), indicating that the remaining 19 molecules may be potential targets for anti-allergic drugs. Analysis of cell

Table 2 Mast cell-, basophil- and eosinophil-selective transcripts for ion channels and receptors

Transcript (accession No. at GenBank)	Cell-type selectivity
Ion channels	
Ca^{2+} channel type A1 D (BE550599)	Basophils, eosinophils
Histamine H_4 receptor (AF312230.1)	Basophils
Prostaglandin E receptor type 3a2 (X83858.1)	Basophils
C3a receptor (U62027.1)	Basophils, eosinophils
Chemokine receptor CCR3 (NM_001837.1)	Basophils, eosinophils
CRTH2 (NM_004778.1)	Basophils, eosinophils
EMR-1 (NM_001974.1)	Basophils, eosinophils
Adenosine A_3 receptor (NM_000677.2)	Eosinophils
$P2Y_2$ purinergic receptor (NM_002564.1)	Eosinophils
GPR105 purinergic receptor (NM_014879.1)	Eosinophils
Other receptors	
FcεRIα (BC005912.1)	Basophils
HTm4 (L35848.1)	Basophils
IL-3 receptor α (NM_002183.1)	Basophils
CD244 NK cell receptor† (NM_016382.1)	Basophils, eosinophils
Fibroblast growth factor receptor 2†(NM_022969.1)	Basophils, eosinophils
IL-5 receptor α (M75914.1)	Basophils, eosinophils
Siglec 8 (NM_014442.1)	Eosinophils
CD117 c-KIT (NM_000222.1)†	Mast cells
Siglec 6 (D86358.1)	Mast cells
FcεRIβ (NM_000139.1)	Mast cells, basophils
Low-density lipoprotein receptor† (NM_000527.2)	Mast cells, basophils
TRK receptor (NM_002529.2)	Mast cell, basophils

Cell type-selective transcripts were chosen from approximately 20 000 transcripts (GeneChip U133A; Affymetrix) based on the following criteria: (i) the average mRNA expression level of each gene (obtained from three independent experiments) in a certain cell type must be threefold or greater than the maximal level in other cell types; (ii) the average mRNA expression level of each gene must be significantly ($P < 0.01$) greater than that in other cell types; (iii) the expression level provided with 'absence' call by GeneChip Software should be observed only once or not at all in three independent experiments; (iv) for transcripts preferentially expressed for the two different cell types, the average normalized expression levels in the two cell types should be within threefold of each other.

CRTH2, chemoattractant receptor, homologous molecule expressed on Th2 cells; EMR, epidermal growth factor-like module-containing mucin-like receptor; TRK, tropomyosin-related kinase neurotrophin; GPR, G-protein-coupled receptor; IL, interleukin.

†Transcripts highly expressed by more than five different normal organs (http://www.lsbm.org/index_e.html).

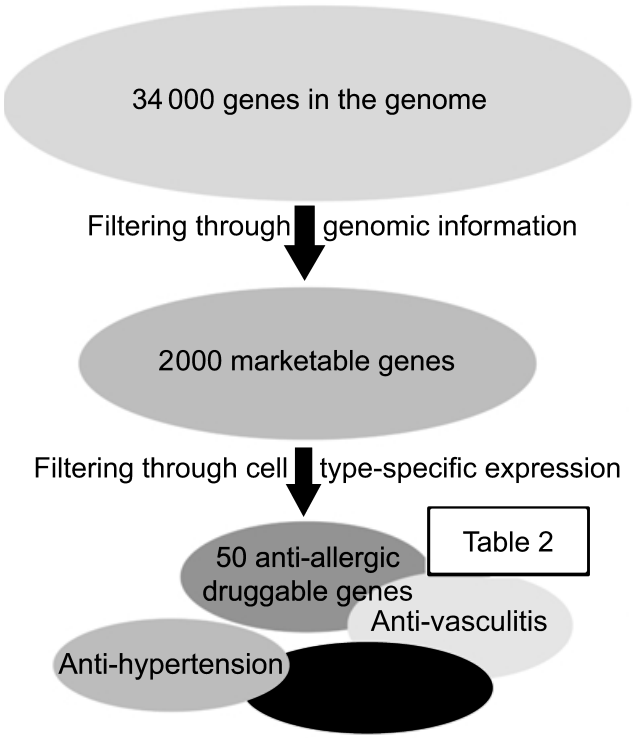

Fig. 2 Filtering of 'druggable' genes through genomic information. Only 34 000 genes were found to be contained in the human genome. Using marketable (= easy to obtain the specific antagonist or agonist) criteria, 2000 genes were filtered as drug targets. Using a cell-type specific transcriptome database, these genes were further subclacified to be anti-allergic drug targets or drug targets for other diseases. The estimated number of anti-allergic (mast cell-, basophil- and eosinophil-specific) drug targets was approximately only 50.

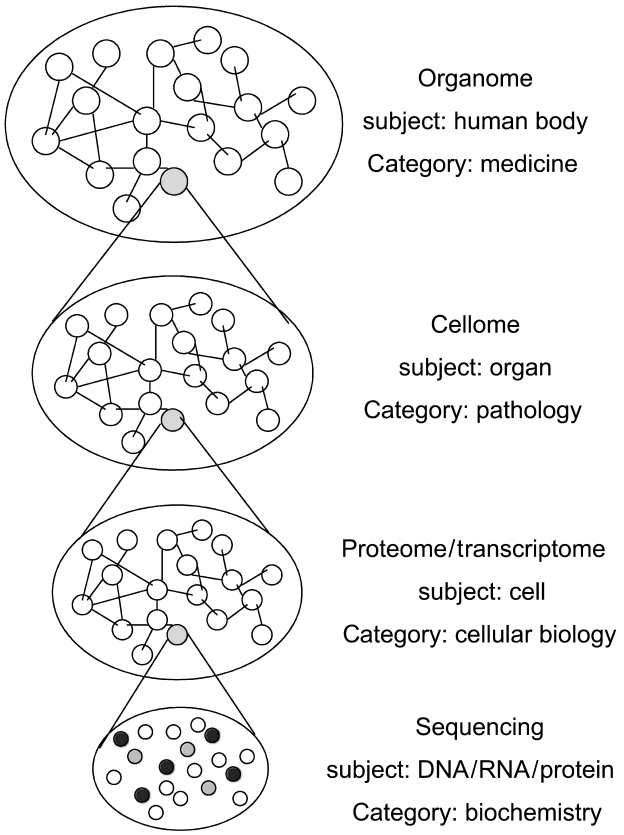

Fig. 3 Stairway to simulation biology. Before establishing a simulation of the human body (organome), which consists of a variety of organs (cellome), which consist of various types of cells (proteome or transcriptome), which consist of various types of proteins (sequencing), which consist of 20 types of amino acids, there are many steps to be understood.

type-selective transcripts from database searches is expected to minimize the efforts required for drug discovery. By using genomic information, we can now filter 2000 marketable genes out of 34 000 genes present in the human genome. Using the comparative expression database regarding each cell type, we will be able to further filter approximately 50 druggable genes for allergic inflammation with regard to safety in addition to marketability (Fig. 2).

Is that all there is, my friend?

When you were a little child, you felt the world or even your town spreading vast or endless. Now, you have found it so small and perhaps you can easily draw the map. Such was the whole of mankind. We once believed that the genome contained almost limitless information. However, after completion of sequencing of the human genome, we, the grown-up mankind, would not think so. We are not trying to discover the 12th Toll-like receptor any more, because we already know that there are no more than 10 or 11 Toll-like receptors by searching the genome structure. In the near future, you may not be able to find a new molecule in the human body. Instead, in the next decade, you will find many new molecule networks present in a cell. Then, you will be able to use computational modeling to analyze integrative biological function and test your hypothesis without performing any *in vivo* or *in vitro* experimentation. However, before establishing a simulation of the human body, which consists of a variety of organs, consisting of various types of cells, consisting of various types of proteins, consisting of 20 types of amino acids, there are many steps that need to be understood (Fig. 3).

REFERENCES

1 Saito H. Translation of the human genome into clinical allergy. *Allergol. Int.* 2003; **52**: 65–70.
2 Hunter PJ, Borg TK. Integration from proteins to organs: The Physiome Project. *Nat. Rev. Mol. Cell Biol.* 2003; **4**: 37–43.
3 Ueda HR, Chen W, Adachi A *et al*. A transcription factor response element for gene expression during circadian night. *Nature* 2002; **418**: 534–9.
4 Nakajima T, Inagaki N, Tanaka H *et al*. Marked increase in CC chemokine gene expression in both human and mouse mast cell transcriptomes following Fcε receptor I cross-linking: An interspecies comparison. *Blood* 2002; **100**: 3861–8.
5 Nakajima T, Iikura M, Okayama Y *et al*. Identification of granulocyte subtype-selective receptors and ion channels by high-density oligonucleotide probe array. *J. Allergy Clin. Immunol.* 2004; **133**: 528–35.
6 Persson CGA. Mice are not a good model of human airway disease. *Am. J. Respir. Crit. Care Med.* 2002; **166**: 6–7.
7 Gelfand EW. Mice are a good model of human airway disease. *Am. J. Respir. Crit. Care Med.* 2002; **166**: 5–6.
8 Saito H, Hatake K, Dvorak AM *et al*. Selective differentiation and proliferation of hematopoietic cells induced by recombinant human interleukins. *Proc. Natl Acad. Sci. USA* 1988; **85**: 2288–92.
9 Shichijo M, Inagaki N, Nakai N *et al*. The effects of anti-asthma drugs on mediator release from cultured human mast cells. *Clin. Exp. Allergy* 1998; **28**: 1228–36.
10 Hamelmann E, Oshiba A, Loader J *et al*. Antiinterleukin-5 antibody prevents airway hyperresponsiveness in a murine model of airway sensitization. *Am. J. Respir. Crit. Care Med.* 1997; **155**: 819–25.
11 Leckie MJ, ten Brinke A, Khan J *et al*. Effects of an interleukin-5 blocking monoclonal antibody on eosinophils, airway hyper-responsiveness, and the late asthmatic response. *Lancet* 2000; **356**: 2144–8.
12 Kwitek AE, Tonellato PJ, Jacob HJ *et al*. Automated construction of high-density comparative maps between rat, human, and mouse. *Genome Res.* 2001; **11**: 1935–43.
13 Nakajima T, Matsumoto K, Suto H *et al*. Gene expression screening of human mast cells and eosinophils using high-density oligonucleotide probe arrays: Abundant expression of major basic protein in mast cells. *Blood* 2001; **98**: 1127–34.
14 Bochner BS. Systemic activation of basophils and eosinophils: Markers and consequences. *J. Allergy Clin. Immunol.* 2000; **106**: S292–302.
15 Savage CO. The evolving pathogenesis of systemic vasculitis. *Clin. Med.* 2002; **2**: 458–64.
16 Oehler R, Weingartmann G, Manhart N *et al*. Polytrauma induces increased expression of pyruvate kinase in neutrophils. *Blood* 2000; **95**: 1086–92.
17 Church MK, Pao GJ, Holgate ST. Characterization of histamine secretion from mechanically dispersed human lung mast cells: Effects of anti-IgE, calcium ionophore A23187, compound 48/80, and basic polypeptides. *J. Immunol.* 1982; **129**: 2116–21.
18 Lin CS, Boltz RC, Blake JT *et al*. Voltage-gated potassium channels regulate calcium-dependent pathways involved in human T lymphocyte activation. *J. Exp. Med.* 1993; **177**: 637–45.
19 Saito H, Okajima F, Molski TF, Sha'afi RI, Ui M, Ishizaka T. Effects of ADP-ribosylation of GTP-binding protein by pertussis toxin on immunoglobulin E-dependent and -independent histamine release from mast cells and basophils. *J. Immunol.* 1987; **138**: 3927–34.
20 Ji TH, Grossmann M, Ji I. G protein-coupled receptors. I. Diversity of receptor–ligand interactions. *J. Biol. Chem.* 1998; **273**: 17 299–302.
21 Zambrowicz BP, Sands AT. Knockouts model the 100 best-selling drugs. Will they model the next 100? *Nat. Rev. Drug Discov.* 2003; **2**: 38–51.
22 Venter JC, Adams MD, Myers EW *et al*. The sequence of the human genome. *Science* 2001; **291**: 1304–51.
23 Drews J. Drug discovery: A historical perspective. *Science* 2000; **287**: 1960–4.
24 Hopkins AL, Groom CR. The druggable genome. *Nat. Rev. Drug Discov.* 2002; **1**: 727–30.