ISSN 1022-7954, Russian Journal of Genetics, 2015, Vol. 51, No. 4, pp. 408–417. © Pleiades Publishing, Inc., 2015.
Original Russian Text © V.P. Puzyrev, 2015, published in Genetika, 2015, Vol. 51, No. 4, pp. 491–502.

Genetic Bases of Human Comorbidity

V. P. Puzyrev

Institute of Medical Genetics, Siberian Branch of the Russian Academy of Medical Sciences, Tomsk, 634050 Russia
e-mail: p.valery@medgenetics.ru
Department of Medical Genetics, Siberian State Medical University, Tomsk, 634050 Russia
National Research Tomsk State University, Tomsk, 634050 Russia

Received October 14, 2014; in final form, November 14, 2014

Abstract—In this review, the development of ideas focused on the phenomenon of disease combination (comorbidity) in humans is discussed. The genetic bases of the three forms of the phenomenon, comorbidity (syntropias), inverse comorbidity (dystropias), and comorbidity of Mendelian and multifactorial diseases, are analyzed. The results of personal genome-wide association studies of the genetic risk profile that may predispose an individual to cardiovascular disease continuum (CDC), including coronary heart disease, type 2 diabetes, hypertension, and hypercholesterolemia (CDC syntropy), as well as the results of bioinformatic analysis of common genes and the networks of molecular interactions for two (bronchial asthma and pulmonary tuberculosis) diseases rarely found in one patient (dystropy), are presented. The importance of the diseasome and network medicine concepts in the study of comorbidity is emphasized. Promising areas in genomic studies of comorbidities for disease classification and the development of personalized medicine are designated.

DOI: 10.1134/S1022795415040092

INTRODUCTION

In studies on the genetics of complex traits (normal and pathological) in humans, along with the undoubted advances in genotyping, the problems of phenotyping, with respect to trait infinity and the lack of a unified approach to their systematics, are increasingly being discussed [1, 2].

In human populations, about 40 million genetic variants including such structural changes as insertions/deletions, copy number variations (CNVs), inversions, and SNPs (which constituting 95% of all known sequence variants) were identified [3]. In genome-wide association studies (GWAS), statistically significant associations with more than 8900 SNPs were demonstrated for almost 300 diseases and traits [4]. At the same time, only 16% of the risk-associated loci are loaded with SNPs affecting coding sequences. These data indicate the nonparticipation of the majority of risk loci in changes in the amino acid sequences of the proteins [5]. An important task of postgenome-wide studies is to analyze the associations of the entire set of genetic variants of risk locus with gene expression, RNA splicing, transcription factor binding, DNA methylation, and histone modifications [3]. Modern technological achievements make it possible to solve these problems.

Human genotyping is associated with quite different problems. The potential complexity is evident already at the cellular level. Specifically, the human interactome includes at least 1000 metabolites and an indefinite number of different proteins and functional RNA molecules. Furthermore, the number of cellular components, which are the interactome nodes, is more than 100 thousand [6]. Another problem is that the genotyping and all variants of omic projects and technologies in association studies are carried out mainly for individuals suffering from one particular (relatively isolated) pathology, whereas in clinical practice there is the problem of multiple diseases (diseases combinations) [7, 8]. At least one third of the population carries more than one disease [9]. In the United States, 80% of the health budget is spent on patients with four or more diseases [10].

The phenomenon of polypathy was first considered by French clinicians in the second half of the 19th century. Ch. Bouchard proposed the concept of arthritismus, which explained the combination of joint diseases, obesity, diabetes, and early atherosclerosis by a common metabolic disorder (braditrophia) and burdened heredity [11]. Further studies in this area were focused on ideas on the landscape of human diseases, which were based on the data of clinical and epidemiological studies and on the mapping of related pathological conditions [12–14]. Now they are supplemented by the results of genomic and bioinformatics research. The phenomenon of comorbidity [7], which refers to multiple diseases in a single patient, is actively investigated from the standpoint of modern genetic concepts of the diseasome [15] and network medicine [6].

This review presents an analysis of the results of genomic studies of the three forms of comorbidity in humans: direct comorbidity (syntropias), inverse comorbidity (dystropias), and comorbidity between

multifactorial (complex) and Mendelian (monogenic) diseases.

COMORBIDITY: FROM SYNTROPIAS TO NETWORK MEDICINE

The term comorbidity (syn., polypathy, multimorbidity) was introduced in 1970 by an American physician and specialist in the field of epidemiology of noncommunicable diseases, A. Feinstein, and was defined as the manifestation of an additional clinical condition that coexisted or arose within the context of a primary disease [7]. This clinical condition can be a disease, a pathological syndrome, pregnancy, a prolonged strict diet, or a complication of drug therapy. Comorbidity is a complex of several (megaforms, conglomerates) diseases that simultaneously exist in individual patients and are observed more often than would be expected for a random distribution.

Studies of the phenomenon of comorbidity have attracted growing interest. Over ten years (from 1990 to 2000), two reviews on this issue were published. In the following decade, 39 reviews appeared, and the International Research Community on Multimorbidity (IRCM) was founded. The *Journal of Comorbidity* (JOC) has been published since 2010 [16].

The epidemiological data on the prevalence of comorbidity among chronic noncommunicable diseases in different populations vary greatly and depend on the parameters of the studied samples (gender, age, clinical polymorphism of the diseases, commitment of the researchers to different classifications of the disease states). However, the comorbidity index is clearly higher with age, especially in women. For instance, the prevalence of chronic comorbid diseases varies from 2.8 in young women to 6.4 in women of senior age [17]. These indices, obtained based on the analysis of clinical data from patient records, were consistent with the data from pathology reports. Specifically, an analysis of more than 3000 autopsies showed that the frequency of comorbidity was 94% in individuals between the ages of 50 and 70; the combination of two or three diseases was the most common, though the coexistence from six to eight diseases in one patient was observed in single cases (up to 2.7%) [18].

Among young patients (18 to 29 years old) with obesity, more than two chronic disease were found in 22% males and 43% females; 75% of individuals with obesity from this Canadian sample had dyslipidemia, arterial hypertension, and type 2 diabetes [19]. Among older patients with arthritis, 50% had arterial hypertension, 20% had cardiovascular disease, and 14% had type 2 diabetes [20].

Syntropias. Comorbidity may be caused by several reasons. Among them are general environmental factors (environmental, social status, life style), the side effects of common disease treatment schemes, etc. However, the molecular origin of these disease combinations (common genes, molecular pathogenesis) deserves special attention [21, 22]. It is for this category of associated diseases, half a century before the appearance of the term comorbidity, German pathologists, M. Pfaundler and L. von Seht [12], established the term syntropy, which is defined as mutual disposition, or the attraction of two or more diseases in the same individual. According to the suggestion of these authors, such combinations were caused not only by the conditions of life and nutrition, but also by internal features of the organism's reactivity, which they associated with concept (which was very popular in their time) of diatheses, i.e., special conditions of the organism that were inherited and characterized by a tendency to develop certain groups of diseases. The important role of genetic factors in the formation of syntropias was previously suggested upon the description of Pfaundler–Hurler syndrome [23]. This disease, which included several congenital abnormalities (multiple abarts, from the German "abart", malformation) was attributed to syntropias. Actually, it was in this work that the term syntropy first appeared [23].

Based on an analysis of more than 30 thousand medical histories of patients with chronic diseases, a tendency of some disease states for joint manifestations in individual patients and their close relatives was later demonstrated [12]. Such conglomerates of diseases, having similar mechanisms of development (pathogenesis) and genealogical concordance, are more common than theoretically expected with random distribution. Because of these characteristics, these diseases make up a special group among the other variants of disease combinations (comorbidity, multimorbidity). In the modern definition, it is emphasized that syntropy is a natural species-specific combination of two or more pathological conditions (nosologies, syndromes) in an individual and his close relatives that is not accidental and has an evolutionary and genetic basis [24]. A nonrandom combination of individual forms of pathology, united by the similarity of the pathogenesis, and genealogical concordance indicate the possibility of the involvement of common susceptibility genes to the development of the single pathological components and the formation of a particular syntropy. The genes that contribute to the development of syntropias were termed as syntropic genes [24]. More strictly, these are the sets of functionally interacting coregulated genes that are distributed across the whole human genome and involved in the biochemical and physiological pathways common for this syntropy.

There are many convincingly clinically proven syntropic diseases, including cardiovascular diseases united into a cardiovascular continuum (CDC) [25]; immuno-mediated diseases (allergic disease and autoimmune diseases) [26, 27]; endocrine diseases, including the combination of diabetes mellitus, autoimmune thyroiditis, and celiac disease [28]; mental diseases, including depressive and bipolar disorders

[29]; and others. The studies of such disease combinations usually provide no evidence for the role of genetic factors in the comorbidity formation, only various proposals on the issue have been made.

In this category of phenotypic studies comes one of the first important investigations based on the analysis of a clinical database including 1.5 million of medical records, both on monogenic and multifactorial diseases. Using the appropriate probabilistic models and correction for possible bias regarding the gender, age, and ethnicity of the patient at the time of the disease onset and without making any assumptions about the heritability and familial cases of the diseases, the authors of the study found correlations for 161 diseases [13]. In addition to the expected and disease combinations known from clinical observations (for example, the association between schizophrenia and bipolar disorder), some surprising data were obtained. For example, a negative correlation between aortic aneurysm and schizophrenia, as well as between breast cancer and bipolar disorder was observed. These data point to the nonrandom occurrence of certain combinations of diseases in human populations.

Diseasome and network medicine. The improvements in genotyping technologies has provided supplementation of the phenotypic information on relationships (combinations) between human diseases with the construction of gene networks with subsequent analysis of gene–phene associations. However, until recently, most of these successful studies were concentrated on a single disease. Networking tools were used to analyze the interaction between genes and a single disease. To date, a conceptual framework for the study of the relationships of all human diseases (the disease phenome) with the complete collection of disease-controlling genes (the disease genome) was developed (elaborated) by providing the formation of global diseasome pattern, combining all of the known gene–disease associations [15]. According to the definition of the authors, a diseasome is a collection of all known gene–disease associations organized in the network of human diseases (human disease network – HDN). The network consists of the hubs in which the diseases are located and the interconnecting ribs, which are represented by common cause-dependent genes. In particular, a cancer cluster includes a number of interrelated diseases with a strong predisposition to cancer. This is natural, since many oncogenes or tumor suppressors are associated with different types of cancer.

We conducted a genome-wide association study of the genetic profile for susceptibility to diseases of the cardiovascular continuum, which we designated as CDC syntropy [11, 24]. This syntropy included the simultaneous presence of four diseases (pathological states) in a patient: coronary heart disease (CHD), type 2 diabetes (T2D), arterial hypertension (AH), and hypercholesterolemia (HCh). A sample in which the patients were particularly loaded with diseases (CDC syntropy) was compared with two other samples, one of which was represented only by patients with CHD and the other included patients with a combination of two diseases, CHD and AH, and no signs of T2D or HCh. In the study, a total of 1400 genetic markers were used, and the study was performed using the My Gene genomic service platform (www.i-gene.ru).

The only CHD phenotype was associated with 14 polymorphic variants, including those related to the *APOB*, *CD226*, *NKX2-5, TLR2*, *DPP6*, *KLRB1*, *VDR*, *SCARB1*, *NEDD4L*, and *SREBF2* genes, and four genetic markers in the intergenic spacer regions (rs12487066, rs7807268, rs10896449, and rs944289). The CHD in combination with the AH phenotype was associated with 13 genetic markers, including those in the *BTNL2*, *EGFR*, *CNTNAP2*, *SCARB1*, and *HNF1A* genes, and six variants in the intergenic spacer regions (rs801114, rs10499194, rs13207033, rs2398162, rs6501455, and rs1160312). The combination of several diseases (syntropy) was associated with 14 markers, including those in the *TAS2R38*, *SEZ6L*, *APOA2*, *KLF7*, *CETP*, *ITGA4*, *RAD54B*, *LDLR*, *LDLR*, and *MTAP* genes, and three markers in the intergenic spacer regions (rs1333048, rs1333049, and rs6501455). Syntropy and the combination of two diseases (CHD and AH) had two common genetic markers (*SEZ6L* rs663048 and rs6501455); the combination of CHD and AH and only CHD had one common marker (*SCARB1* rs4765623). Syntropy and CHD had no common genes among those studied. The classification analysis of the assignment of associated genes to a particular metabolic pathway showed that lipid metabolism genes were involved in the formation of all three variants (different combinations) of the cardiovascular disease continuum, while the immune response genes were specific to CHD and were not involved in the syntropy formation. This study demonstrated that the genetic profile of the combination of several diseases could be considerably different from individual, non-associated forms of pathology.

A description of the functions of associated genes was supplemented with the classification network analysis of intergenic interactions, which made it possible to trace the circuit of the interactions of several genes. This is STRING analysis (Fig. 1) [http://string-db.org/], which makes it possible to formally correlate a particular gene with the most important metabolic pathways. The use of such formalized approach showed that, among the CDH genes, genes associated with the functioning of the immune system and lipid metabolism dominated. In the case of CHD in combination with the AH phenotype, two genes belonged to the immune system and another two belonged to the lipid metabolism. Namely, *SCARB1*, the lipid metabolism gene, was common to these two forms of pathology. Among the syntropy-associated genes, three genes were associated with lipid metabolism. In the STRING analysis, other genes were attributed to any

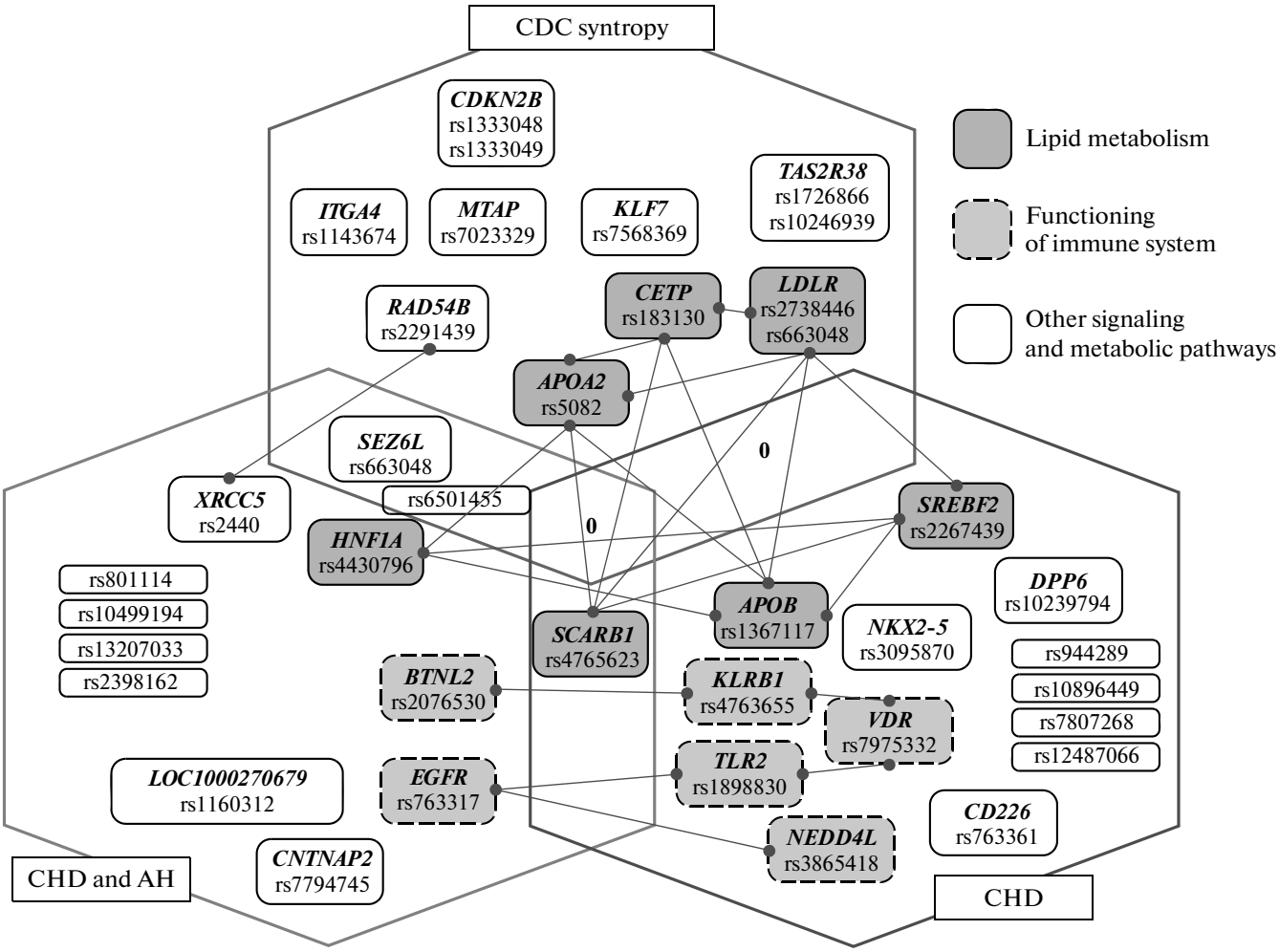

Fig. 1. Venn diagram of relationships among the genes of groups with intergenic interactions. Intergenic relationships were designed based on data from the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) online service [http://string-db.org].

metabolic pathway. Thus, our data suggested that the lipid metabolism genes were involved in all variants of the disease course from the cardiovascular disease continuum (including their different combinations), while the immune system regulator genes were specific to CHD and were not involved in the formation of CDC syntropy.

In 2008, Lee et al. [30] created a scheme of the metabolic disease network, in which the ribs were placed between hubs and diseases if the defective enzyme associated with the diseases was responsible for contiguous reactions in metabolic pathways. It was demonstrated that pairs of associated diseases exhibited considerable similarity in the expression patterns of the genes encoding the corresponding enzymes. The network topological approach unravels new mechanisms of a single disease on the basis of the common pathophysiology of a conglomerate of associated diseases. The authors pointed to the the very important detail that comorbidity of metabolic disorders could be better predicted by assessing the links between disrupted metabolic reactions, as compared the predictions based on the existence of common disease-responsible genes. This was noted by Piro [22], who analyzed works that focused on network medicine. He showed that the network density of comorbid diseases was noticeably higher than the network density of genetic diseases. These findings indicated that a high level of comorbidity may exist, despite the absence of common genes [22].

The construction of networks underlying the biological processes and metabolic pathways associated with rare (orphan) diseases (OD), which are those characterized by frequencies of less than 65 per 10 thousand inhabitants, seems to be reasonable [22]. The first orphan disease network resulted from research by Zhang et al. [31], who discovered the unexpected pathophysiological and genetic relationships of three disease categories: monogenic orphan diseases, orphan diseases for which there were several causative

genes, and complex diseases with genetic predisposition. It was found that the mutant genes causing orphan diseases mostly encode essential proteins, whereas most of the genes for common widespread diseases are not essential and do not encode hub proteins. Approximately 62% of orphan disease genes are essential genes; their knockout in mice is lethal. It is noticeably higher than the 18% of essential genes in the network of widespread diseases of a multifactorial nature.

It is noteworthy that the functional relationship between the causative genes involved in different OD, when taken into account, allows the identification of a common molecular bases of pathogenesis of these diseases. Such associations may be used in the future development of new hypotheses on the molecular mechanisms of disease development, to substantiate relevant therapy, and to identify potential candidates for drug repositioning [32]. A global analysis of all OD can facilitate the analysis of the causes of comorbidity [31].

INVERSE COMORBIDITY (DYSTROPIAS)

The relationships between chronic diseases in a particular individual (the combination profile and synchrony, the time of their simultaneous onset) is not limited to the phenomenon of direct (positive) comorbidity. In parallel to the concept of syntropy, as a variant of comorbidity, the term dystropy (repulsion) was suggested for those pathologies that were rarely found in the same patient at the same time [12]. Such antagonistic relationships between diseases (counterassociations) were termed as inverse comorbidity [33], which is essentially a synonym for the old term, dystropy.

There is epidemiological evidence that patients with Down's syndrome, Parkinson's disease, schizophrenia, diabetes mellitus, anorexia nervosa, Alzheimer's disease, multiple sclerosis, and Huntington's disease are protected from many forms of cancer, including solid tumors, tumors associated with smoking, and the development of prostate cancer [33]. These relationships at a clinical level were also described for other disease categories. Such dystropias as tuberculosis and bronchial asthma, type 1 diabetes and peptic ulcer, proliferative processes of lymphoid and myeloid types, and single cases of the simultaneous development of B and T cell lymphomas in a single individual have been described [11, 34].

The disease associations at the level of clinical phenotypes (across the nosological boundaries) have a molecular genetic basis, i.e., common genes and overlapping metabolic pathways. However, the paradoxical nature of inverse comorbidity was mentioned, meaning that dystropic genes appeared to be the same for diseases as their counterassociative relationships, but they exhibited differently directed expression [33].

New molecular evidence of inverse comorbidity between central nervous system disorders and cancer were recently obtained by means of transcriptomic meta-analysis [35]. The authors conducted transcriptomic analysis of three central nervous system (CNS) pathologies (Alzheimer's disease, Parkinson's disease, and schizophrenia) and three types of cancer (lung, prostate, rectum), for which inverse comorbidity was previously described. This approach enables the identification of the candidate genes potentially associated with inverse comorbidity. For instance, it was demonstrated that 74 genes were simultaneously suppressed in three CNS disorders, while their activity was increased in the three investigated types of cancer. On the contrary, the expression of 19 genes was simultaneously increased in the three studied CNS pathologies, while it was suppressed in the three types of cancer. Moreover, a comparison of differentially expressed genes of the two groups of diseases (CNS and cancer) with other diseases for which no inverse comorbidity was described (asthma, HIV, malaria, sarcoidosis) showed no differently directed changes in the regulation of gene expression (characteristic of CNS diseases and cancer). Discussing these results, the authors made a suggestion that could have innovative value. The postmortem brain samples (in the case of CNS disorders) or tumor tissues (in the case of cancer) were probably obtained from patients who received medication. It can be thus suggested that the observed changes in the regulation of gene expression could be due to the action of these drugs. In this case, it can be assumed that some of the drugs used to treat CNS disorders can cause the reversion of the expression of some genes that control the development of cancer. Further research in this field may open new directions in the search for effective drugs.

We carried out an analysis of common genes for susceptibility to bronchial asthma (BA) and pulmonary tuberculosis (TB), the two diseases that, according to epidemiological data, are rarely observed in one individual [36]. With the molecular interaction network reconstructed with the help of the ANDSystem software program, it was demonstrated that asthma and tuberculosis are closely related to each other, as compared with 10 thousand randomly selected disease pairs. The majority of BA and TB common genes belongs to the cytokine-coding genes (interleukins, interferons, tumor necrosis factor, chemokines). For four of these genes, *IL2*, *IL10*, *IL12B*, and *VDR*, differently directed and opposite effects (risk increasing and protective manifestations) of one and the same genotype, relative to the development of BA and TB, were demonstrated. A database search for associations of these genes with BA and TB showed that the *IL2* T allele (rs 2069762) contributes to the risk of BA development [38]. At the same time, a protective role of this allele for TB was reported [39]. Similarly, the combination of GA + AA genotypes (rs 1800896) of the *IL10* gene provided a protective effect on the

development of TB (OR = 0.55; 95%CI, 0.35–0.88; $p = 0.01$), while the AA genotype was associated with an increased risk of BA (OR = 1,26; 95%CI, 1.02–1.55; $p < 0.05$) [40, 41]. Analysis of the rs 228570 polymorphic variant of the *VDR* gene showed that the ff genotype was associated with the risk of TB (OR = 1.91; 95%CI, 1.44–2.52) [42], and the protective value of this genotype relative to BA was indicated [43]. Genotype 2/2 of the *IL12B* rs3212227 polymorphism reduced the risk of TB [44] and contributed to the development of BA [45].

Thus, inverse comorbidity provides an unprecedented opportunity to clarify the pathogenesis of many widespread and socially important diseases, as well as the understanding of why some individuals with a diagnosis of specific disorders are protected from other diseases. This information is important for the new methods of treatment [33].

COMORBIDITY BETWEEN MULTIFACTORIAL AND MONOGENIC DISEASES

A recent analysis of more than 110 million electronic medical records provided a new understanding of the relationships of monogenic (Mendelian) diseases (MD) and multifactorial (complex) diseases (MFD) [46]. A total of 2909 paired comorbidities between MFDs and MDs and 462 such associations between MDs were revealed. The presence of these associations is not surprising, though their wide distribution was unexpected. In a study based on an analysis of a number of databases, it was demonstrated that 54% (524 out of 968) of the genes with mutations that cause Mendelian diseases were also involved in the occurrence and development of MFDs [47].

An analysis of millions of electronic clinical records obtained from different regions of the United States and Denmark performed by D.R. Blair et al. [46] was based on the hypothesis of transitive association. According to this hypothesis, in the case of comorbidity between multifactorial and Mendelian diseases, the risk of MFD is determined among others by the genes causing risk-associated MD.

Along with the well-known cases of comorbidity (for example, comorbidity between lipoprotein insufficiency and myocardial infarction, ataxia–telangiectasia and breast cancer), a phenotypic comorbidity was discovered between Marfan syndrome and several neuropsychiatric diseases (autism, bipolar disorder, depression) and between X-chromosome fragility and asthma, psoriasis, and viral infection. This reflects potential immune system dysfunction among these patients [48]. The authors suggested that the comorbidity between MFDs and MDs indicated that the genes controlling the MDs can make nonadditive contributions to MFD risk and induced various pathological consequences in accordance with the Mendelian code. Common risk variants associated with MFD are often concentrated in the comorbid Mendelian loci. In this sense, the patients included in genome-wide association studies (GWAS) carry genetic variants predisposed to both MD and MFD. This hypothesis is confirmed by several examples.

For example, according to GWAS results, four out of seven tumors were associated with both rare and common variants of the *TERT* locus, which encodes human telomerase reverse transcriptase. Some of these variants completely inhibit enzyme activity and lead to dyskeratosis congenita syndrome [49]. Rare mutations in the promoter region of the gene associated with the familial melanoma were recently identified, and their carriers were also characterized by an increased risk of the development of other tumors [50]. The analysis of comorbidity performed in this study showed that schizophrenia, bipolar disorder, autism, and depression could be associated with the *SYNE1*, *PRPF3*, *CACNA1C*, and *PPP2R2B* Mendelian loci. These loci were also found to contain frequent polymorphisms associated with the risk of the same diseases. The authors refer to the results of exome sequencing performed in patients with autism and identified both inherited and *de novo* mutations in the *SYNE1* [51, 52] gene. These data suggested that the examined genes could also have some other rare variants that caused a predisposition to many neuropsychiatric disorders in their carriers. If this is true, then the strategy of combining genomic sequencing data in patients with these different albeit related phenotypes of multifactorial nature can be a way to improve the efficiency of identifying rare gene variants manifesting moderate effects.

Thus, in accordance with the hypothesis of transitive associations, it can be assumed that each MFD has a unique architecture of the Mendelian disease alleles, creating the so-called nondegenerate code, in which the probability of each disease associated with its Mendelian loci is recorded. Analysis of the MFD–MD comorbidity can be used to detect the Mendelian loci holding the genetic variants of predisposition to MFD. The genes responsible for MD carry both rare and frequent adverse variants, and alleles of the whole spectrum contribute to the disease risk (the allelic series hypothesis). This opinion is commonly held by other researchers [53]. The GWAS design enables the identification only frequent variants; rare variants are usually detected through linkage analysis and sequencing. Investigation of the MFD–MD comorbidity represents an additional approach to finding and assessing the role of rare genes in the MFD pathogenetics.

INTERSPECIFIC PHENOTYPIC ASSOCIATIONS: PHENOLOGS AND NEW GENES OF HUMAN DISEASES

Studies based on the method of constructing simultaneously analyzed disease networks and gene networks confirmed the existence of global phenome

Some examples from the more than 6200 phenologs identified upon the comparison of human diseases (Hs) and mutant phenotypes in mouse (Mm), yeast (Sc), and *Arabidopsis thaliana* (At) (according to [54])

Species 1	Phenotype 1	Species 2	Phenotype 2	$n1$	$n2$	k	p
Hs	X-linked conductive hearing loss	Mm	Animal circling	47	50	12	2×10^{-20}
Hs	Bardet–Biedl syndrome	Mm	Tailless sperm	11	5	4	8×10^{-13}
Hs	Zellweger syndrome	Sc	Reduction in the number of peroxisomes	8	6	4	1×10^{-9}
Hs	Predisposition to autism	Mm	Abnormal social behavior	5	16	3	1×10^{-8}
Hs	Refsum disease	At	Defects in peroxisomal matrix protein import	4	5	2	1×10^{-5}
Hs	Mental retardation	At	Defects in cotyledon development	13	5	2	1×10^{-4}
Hs	Hemolytic anemia	Sc	Hydroxyurea sensitivity	11	23	3	2×10^{-4}
Hs	Amyotrophic lateral sclerosis	Sc	Increased wortmannin resistance	2	34	2	2×10^{-4}

$n1$, the number of orthologs in species 1 with phenotype 1; $n2$, the number of orthologs in species 2 with phenotype 2; k, the number of overlapping orthologs in the two species.

organization. However, based solely on human diseases with a known molecular basis and with known genes, this approach restricts the possibility of discovering new mechanisms and the prediction of new genes responsible for the disease development [22]. This limitation is reduced by a conceptual expansion of the search space from genes to phenotypes (diseases) in different species. This was done in the study of K.L. McCary et al. [54], where a considerable number of interspecific phenotypic associations, quite obscure and surprising from the physiological positions, were found. For example, retinoblastoma in humans and ectopic nematode vulva are the result of a defect in the retinoblastoma 1 gene in human and in its ortholog in the nematode. The authors of the study used the new approach, enabling quantitative and systematic identification of an unobvious equivalence (parity) between the phenotypes of different species, based on overlapping sets of orthologous genes in human, mouse, yeast, worms, and plants.

The principle of the method is as follows. The orthologs, i.e., two genes diverged from a common ancestor (from one ancestral gene), determine the same character in two or more modern species. Interspecific phenotypes determined by orthologous genes are termed as orthologous phenotypes or phenologs. Phenologs display amazing evolutionary preservation (conservation) of their determining gene networks. If the ortholog overlap in two species is statistically significant, then there can be genes among the group of such orthologs in the two compared organisms that make up the genetic basis of similar characters. On this basis, it is possible to predict even completely new genotype–phenotype associations. Examples of phenologs identified by a comparison of human diseases and mutant phenotypes in model organisms are demonstrated in the table. The yeast model is suggested for detecting the defects of angiogenesis; the worm model is suggested for breast cancer; the mouse model is suggested for autism; and the plant model is suggested for the development of the neural crest defect associated with Waardenburg syndrome [53].

The authors who studied these models observed a striking phenolog between humans and plants; it associated with the defect of negative gravitropism and Waandenburg syndrome. The latter is a congenital syndrome that results from abnormal neural crest development and is characterized by craniofacial dysmorphia, abnormal pigmentation, and hearing loss (in fact, it accounts for up to 5% of all cases of deafness in humans). This phenolog is associated with three genes involved in directed plant growth in response to the force of gravity (gravitation). It was suggested that these genes could be involved in the formation of directed migration of neural crest cells and differentiation during early embryonic development in females. This assumption had some support. It was known that one of the identified proteins (STX12) in mice was involved in pigmentation and hearing defects. Two other proteins were not described in literature. However, with respect to one of them, the sec23ip protein, which was detected based on the identification of the ortholog responsible for the gravitropism defect in plants, it was demonstrated that its expression, evaluated using targeted microinjection of morpholino (MO), led to a noticeable defect in neural crest cell migration at the site of injection, proving the role of the *SEC23IP* gene in the development of neural crest

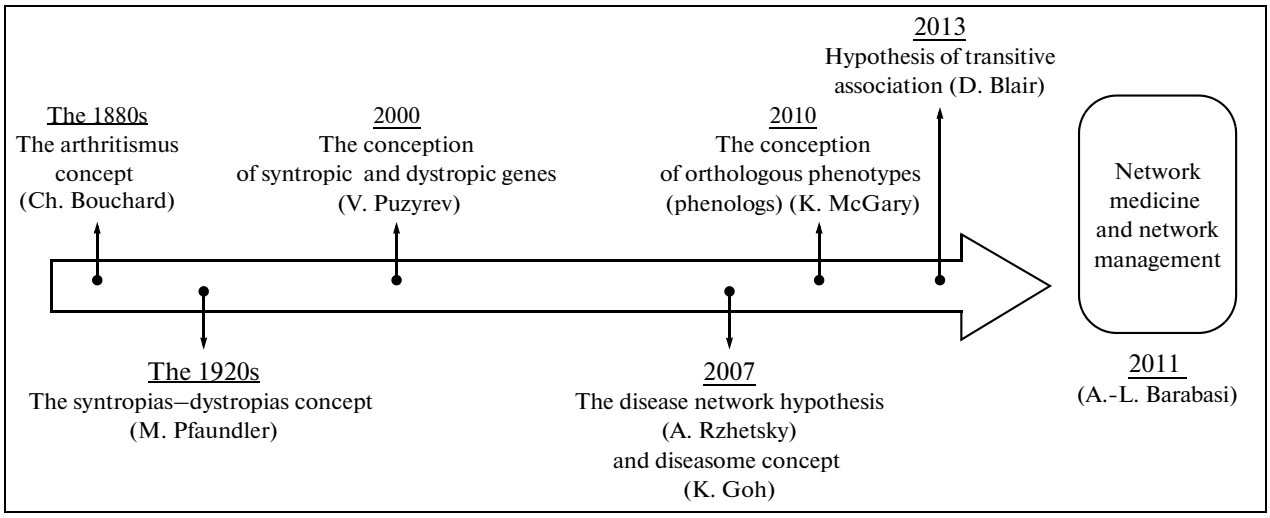

Fig. 2. Stages of the comorbidity analysis in humans.

cells. In the opinion of the authors, *SEC23IP* was a promising candidate gene for genetically heterogeneous Waardenburg syndrome.

In addition, the value of this approach, along with the discovery of new genes of human disease, seems to be important in the revision of disease classifications and in developing methods of diagnosis [14]. In fact, this is the way for a radical increase in the number of gene–phenotype associations in the disease landscape in humans, as well as in model organisms. It is suggested that phenologs can be used for predicting the genes associated with one third or even a half of tested human diseases [54].

INNOVATIVE EXPECTATIONS

Widely known in the clinic and yet remaining enigmatic, the phenomenon of comorbidity, as it is studied by means of genomic and network approaches, is unfolding more fully, giving rise to new concepts and offering new classifications, as well as prospects for the application of new knowledge in genomic medicine. The important stages of this pathway are demonstrated in Fig. 2. It should be emphasized that we are approaching the stage of possible practical application of the theoretical constructions (diseasomes) in the network management (network medicine) [6, 55]. The author of the network medicine concept, one of the pioneers of network biology, A.-L. Barabasi, designated the network medicine paradigm as "thinking globally, acting locally" [6].

An important aspect of thinking about the application of network medicine to the comorbidity challenges is the problem of the treatment of these diseases. First, it should be noted that comorbidity, being common in the practice of the modern physician, is often accompanied by polypragmasy, i.e. the prescription of multiple remedies simultaneously in the effort to cure all diseases constituting a particular syntropy. This often becomes dangerous, causing drug side complications (iatrogeny). In this regard, it seems reasonable to find alternative approaches to the treatment of multiple comorbidities.

One of these approaches is hub therapy for syntropic comorbid diseases; this therapy is aimed at the modulation or even disintegration of the hub networks simultaneously involved in the regulation of several signaling pathways that are common to the corresponding syntropy [16]. The authors of this approach, based on experimental data showing that the removal of 5–15% of hubs led to network disintegration, demonstrated that for patients with early forms of coronary atherosclerosis in combination with autoimmune diseases (rheumatoid arthritis, psoriasis), statins were common, effective, and safe drugs.

It is assumed that the data obtained on the diseasome of orphan diseases (OBs), for which search and development of drugs is considered unprofitable due to rarity of this group of diseases, can make a difference. Studying the diseasome of OBs, especially in the construction of networks of functional interactions of OBs rather than gene ones only may be useful for predicting the development of diseases, as well as for identification of genetic modifiers or drug targets specific for OBs (similar to other "-ome" projects, this approach is called "drugome") [31].

It is appropriate to note that the progress in genomics strengthened the position of personalized medicine. Indeed, medical practice already has many examples of so-called precision medicine aimed at therapeutic correction of specific gene mutations [56]. At the same time, development of the genetic basis of MD comorbidity, as well as of MFD comorbidity, opens prospects for creating identical treatment schemes for diverse diseases with common network

nodes. And such a universal medicine should coexist with personalized medicine.

In addition to the practical benefits to the disclosure of the nature of specific diseases and the development of new therapeutic strategies, an analysis of comorbidity, phenotype networks, and gene networks (diseasome analysis) that deepens our knowledge of the topology of the human phenome can stimulate the revision of current disease classification [14, 22] and of the isolation in these classifications of disease subtypes with a different prognosis for the patient and family members, with an understanding of the differential response to treatment [57].

Finally, it should be emphasized, that complex interactions between the genes, proteins, or cellular pathways formed in the conditions of comorbidity appear in a completely different manner than that for individual diseases treated as independent events. However, for both cases (comorbidity and one specific disease) in modern studies on disease networks, the used information does not always reflect the essence of the dynamic nature of biological systems. The protein interactome data, for example, are often static, although the protein–protein interactions can be of no importance if the protein-coding genes are not expressed in the examined tissues [58]. Therefore, the recommendations for future research to shift the emphasis from static to what is referred to as dynamic network medicine [22] seems to be reasonable.

ACKNOWLEDGMENTS

This study was supported by the Grant of the Russian Science Foundation: Conduction of Fundamental Research and Exploratory Research by Individual Research Groups (grant no. 14-15-00305).

REFERENCES

1. Venter, J.C., Multiple personal genomes await, *Nature*, 2010, vol. 464, pp. 676–677.
2. Puzyrev, V.P. and Kucher, A.N., Evolutionary ontogenetic aspects of pathogenetics of chronic human diseases, *Russ. J. Genet.*, 2011, vol. 47, no. 12, pp. 1395–1405.
3. Zhang, X., Bailey, S.D., and Lupien, M., Laying a solid foundation for Manhattan—"setting the functional basis for the post-GWAS era", *TIG*, 2014, vol. 30, no. 4, pp. 140–149.
4. Hindorff, L.A., MacArthur, J., Morales, J., et al., A Catalog of Published Genome-Wide Associations Studies. http://www.genome.gov/gwastudies
5. Schaub, M.A., Boyle, A.P., Kundaje, A., et al., Linking disease association with regulatory information in human genome, *Genome Res.*, 2012, vol. 22, pp. 1748–1759.
6. Barabasi, A.-L., Gulbahce, N., and Loscalzo, J., Network medicine: a network-based approach to human disease, *Nat. Rev. Genet.*, 2011, vol. 12, no. 1, pp. 56–68.
7. Feinstein, A.R., The pre-therapeutic classification of comorbidity in chronic diseases, *J. Chronic Dis.*, 1970, vol. 23, no. 7, pp. 455–468.
8. Valderas, J.M., Increasing clinical, community and patient-centered health research, *J. Comorbidity*, 2013, vol. 3, no. 2, pp. 41–44.
9. Nikolaev, Yu.A., Mitrofanov, I.M., Pospelov, T.I., et al., Comorbidity in modern internal medicine, *Byull. Sib. Otd. Russ. Akad. Med. Nauk*, 2014, vol. 34, pp. 44–48.
10. Valderas, J.M., Starfield, B., Sibbald, B., et al., Defining comorbidity: implications for the understanding of health and health services, *Ann. Fam. Med.*, 2009, no. 7, pp. 357–363.
11. Puzyrev, V.P., Makeeva, O.A., and Freidin, M.B., Syntropy, genetics testing and personalized medicine, *Person. Med.*, 2010, vol. 4, no. 7, pp. 399–405.
12. Pfaundler, M. and von Seht, L., Über Syntropie von Krankheitszuständen, *Z. Kinderheilk.*, 1921, vol. 30, pp. 298–313.
13. Rzhetsky, A., Wajngurt, D., Park, N., and Zheng, T., Probing gentic overlag among complex human phenotypes, *Proc. Natl. Acad. Sci. U.S.A.*, 2007, vol. 104, pp. 11694–11699.
14. Linghu, B., Snitkin, E.S., Hu, Z., et al., Genome-wide prioritization of disease genes and identification of disease–disease association from an integrated human functional linkage network, *Genome Biol.*, 2009, vol. 10, no. 9, R. 91.
15. Goh, K.I., Cusich, M.E., Vall, D., et al., The human disease network, *Proc. Natl. Acad. Sci. U.S.A.*, 2007, vol. 104, pp. 8685–8690.
16. Shirinskii, V.S. and Shirinskii, I.V., Comorbid diseases—actual challenge in clinic medicine, *Sib. Med. Zh.*, 2014, vol. 29, no. 1, pp. 7–11.
17. Fortin, M., Bravo, G., Hudon, C., et al., Prevalence of multimorbidity among adult seen in family practice, *Ann. Fam. Med.*, 2005, no. 3, pp. 223–228.
18. Vertkin, A.L. and Skotnikov, A.S., The role of chronic allergic inflammation in the pathogenesis of asthma and its rational pharmacotherapy in patients with comorbidities, *Lech. Vrach*, 2009, no. 4, pp. 61–67.
19. Bruce, S.G., Riediger, N.D., Zacharias, J.M., and Young, T.K., Obesity and obesity-related comorbidities in Canadian first nation population, *Prev. Chronic Dis.,* 2011, vol. 8, no. 1, pp. 1–8.
20. Caughey, G.E., Ramsay, E.N., Vitry, A.J., et al., Comorbid chronic disease, discordant impact on mortality in older people: a 14-year longitudinal population study, *J. Epidemiol. Community Health*, 2010, vol. 64, no. 12, pp. 1036–1042.
21. Vidal, M., Cusick, M.E., and Barabasi, A.-L., Interactome networks and human disease, *Cell*, 2011, vol. 144, pp. 986–998.
22. Piro, R.M., Network medicine: linking disorders, *Hum. Genet.*, 2012, vol. 131, pp. 1811–1820.
23. Hurler, G., Über einen Typ multipler Abartungen, vorwiegend am Skelettsystem, *Z. Kinderheilk.,* 1919, vol. 24, pp. 220–224.
24. Puzyrev, V.P. and Freidin, M.B., Genetic view on the phenomenon of combined diseases in man, *Acta Nat.*, 2009, vol. 3, pp. 52–57.
25. Dzau, V.J., Autman, E.M., Black, H.R., et al., The Cardiovascular Disease continuum validated: clinical abidance of improved patient outcomes. Part I: pathophysiology and clinical trial evidence (risk factors through stable coronary artery disease), *Circulation*, 2006, vol. 114, pp. 2850–2870.

26. Cookson, W., The immunogenetics of asthma and eczema: a new focus on the epithelium, *Nat. Rev. Immunol.*, 2004, vol. 4, pp. 978–988.

27. Zhernakova, A., van Diemen, C.C., and Wiymenda, C., Detecting shared pathogenesis from shared genetics of immune-related diseases, *Nat. Rev. Genet.,* 2009, vol. 10, no. 1, pp. 43–55.

28. Doolan, A., Donaghue, K., Fairchild, J., et al., Use of HLA typing in diagnosing celiac disease in patients with type 1 diabetes, *Diabetes Care*, 2005, vol. 28, pp. 806–809.

29. Harvey, M., Bellean, P., and Barden, N., Gene interactions in depression: pathways out of darkness, *Trends Genet.*, 2007, vol. 23, pp. 547–556.

30. Lee, D.S., Park, K.A., Christakis, Z.H., and Barabasi, A.L., The implication of human metabolic network topology for disease comorbidity, *Proc. Natl. Acad. Sci. U.S.A.*, 2008, vol. 105, no. 29, pp. 9880–9885.

31. Zhang, M., Zhu, Ch., Jacomy, A., et al., The orphan disease networks, *Am. J. Hum. Genet.*, 2011, vol. 88, no. 10, pp. 755–766.

32. Linghu, B. and DeLisi, Ch., Phenotypic connections in surprising places, *Genome Biol.*, 2010, vol. 11, p. 116.

33. Tabares-Seisdedos, R. and Valderas, J.M., Inverse comorbidity: the power of paradox in the advancement of science, *J. Comorbidity*, 2013, vol. 3, pp. 1–3.

34. Krylov, A.A., On the problem of comorbidity of diseases, *Klin. Med.*, 2000, vol. 78, pp. 56–58.

35. Ibanez, K., Boullosa, C., Tabares-Seisdedos, R., et al., Molecular evidence for the inverse comorbidity between central nervous system disorders and cancers defected by transcriptomic meta-analyses, *PLoS Genet.*, 2014, vol. 10, no. 2. e1004173

36. Bragina, E.Yu., Tiys, E.S., Freidin, M.B., et al., Insights into pathophysiology of dystrophy through the analysis of gene networks: an example of bronchial asthma and tuberculosis, *Immunogenetics*, 2014, vol. 66, nos. 7–8, pp. 457–465.

37. Demenkov, P., Ivanisenko, T., Kolchanov, N., and Ivanisenko, V., ANDVisio: a new tool for graphic visualization and analysis of literature mined associative gene networks in the ANDSystem, *In Silico Biol.*, 2011, no. 11, pp. 149–161.

38. Christensen, U., Haagerup, A., Binderup, H.G., et al., Family based association analysis of the IL2 and IL15 genes in allergic disorders, *Eur. J. Hum. Genet.*, 2006, vol. 14, pp. 227–235.

39. Selvaraj, P., Alagarasu, K., Harishankar, M., et al., Cytokine gene polymorphisms and cytokine levels in pulmonary tuberculosis, *Cytokine*, 2008, vol. 43, no. 1, pp. 26–33.

40. Zhang, Y., Zhang, J., Tian, C., et al., The-308 G/A polymorphisms in TNF-α gene is associated with asthma risk: an update by meta-analysis, *J. Clin. Immunol.*, 2011, vol. 31, no. 2, pp. 174–185.

41. Zheng, X.Y., Guan, W.J., Mao, C., et al., Interleukin-10 promoter 1082/–819/–592 polymorphisms are associated with asthma susceptibility in Asians and atopic asthma: a meta-analysis, *Lung*, 2013. doi 10.1007/s00408-013-9519-8

42. Cao, S., Luo, P.F., Li, W., et al., Vitamin D receptor genetic polymorphisms and tuberculosis among Chinese Han ethnic group, *Chin. Med. J.* 2012, vol. 125, no. 5, pp. 920–925.

43. Maalmi, H., Sassi, F.H., Berraies, A., et al., Association of vitamin D receptor gene polymorphisms with susceptibility to asthma in Tunisian children: a case control study, *Hum. Immunol.*, 2013, vol. 2, no. 2, pp. 234–240.

44. Tso, H.W., Lau, Y.L., Tam, C.M., et al., Associations between IL12B polymorphisms and tuberculosis in the Hong Kong Chinese population, *J. Infect. Dis.*, 2004, vol. 190, no. 5, pp. 913–919.

45. Hirota, T., Suzuki, Y., Hasegawa, K., et al., Functional haplotypes of IL-12B are associated with childhood atopic asthma, *J. Allergy Clin. Immunol.*, 2005, vol. 116, no. 4, pp. 789–795.

46. Blair, D.R., Lyttle, Ch.S., Mortensen, J.M., et al., A nondegenerative code of deleterious variant in Mendelian loci contributes to complex disease risk, *Cell*, 2013, vol. 155, pp. 70–80.

47. Jin, W., Qin, P., Jin, L., and Xu, S., A systematic characterization of genes underlying both complex and Mendelian diseases, *Hum. Mol. Genet.*, 2012, vol. 21, no. 7, pp. 1611–1624.

48. Ashwood, P., Nguyen, D.V., Hessl, D., et al., Plasma cytokine profiles in Fragile X subjects: is there a role for cytokines in the pathogenesis?, *Brain Behav. Immun.,* 2010, vol. 24, pp. 898–902.

49. Kirwan, M. and Dokal, I., Diskeratosis congenital, stem cells and telomeres, *Biochim. Biophys. Acta*, 2009, no. 1792, pp. 371–379.

50. Huang, F.W., Hodis, E., Xu, M.J., et al., Highly recurrent TERT promoter mutation in human melanoma, *Science*, 2013, no. 339, pp. 957–959.

51. O'Roak, B.J., Deriziotis, P., Lee, C., et al., Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations, *Nat. Genet.*, 2011, vol. 43, pp. 585–589.

52. Yu, T.W., Chahrour, M.H., Coulter, M.E., et al., Using whole-exame sequencing to identify inherited causes of autism, *Neuron*, 2013, vol. 77, pp. 259–273.

53. Lupski, J.R., Belmont, J.W., Bocrwinkle, E., and Gibbs, R.A., Clan genomics and the complex architecture of human disease, *Cell*, 2011, vol. 147, pp. 32–43.

54. McCary, K.L., Park, T.J., Woods, J.O., et al., Systematic discovery of nonobvious human disease models through orthologous phenotypes, *Proc. Natl. Acad. Sci. U.S.A.*, 2010, vol. 107, no. 14, pp. 6544–6549.

55. Liu, Y.-Y., Slotine, J.-J., and Barabasi, A.-L., Controllability of complex networks, *Nature*, 2011, vol. 473, pp. 167–173.

56. Ledford, H., Drug bests cystic-fibrosis mutation, *Nature*, 2012, vol. 482, p. 145.

57. Manolio, T.A., Bringin genome-wide association findings into clinical use, *Nat. Rev. Genet.,* 2013, vol. 4, no. 8, pp. 549–558.

58. Przytycka, T.M., Singh, M., and Slonim, D.K., Toward the dynamic interactome: it's about time, *Brief. Bioinform.*, 2010, no. 11, pp. 15–19.

Translated by N. Maleeva