THE ETIOLOGY OF MENTAL DISORDERS

Concise, Clear and Synoptical

Ladislav Hosák
Mohammad Malekirad
Klára Látalová

KAROLINUM

The Etiology of Mental Disorders
Concise, Clear and Synoptical

Ladislav Hosák, Mohammad Malekirad, Klára Látalová

———

Reviewed by:
prof. MUDr. Michal Hrdlička, CSc.
prof. MUDr. Tomáš Kašpárek, Ph.D.

Published by Karolinum Press
Prague 2022
Edited by Nathalie Weatherald
Cover and layout by Zdeněk Ziegler
Typeset by DTP Karolinum Press
First edition

ISBN 978-80-246-5147-7
ISBN 978-80-246-5240-5 (online : pdf)

Univerzita Karlova
Nakladatelství Karolinum

www.karolinum.cz
ebooks@karolinum.cz

Authors

Prof. MUDr. Ladislav Hosák, Ph.D. – Department of Psychiatry, Charles University, School of Medicine in Hradec Králové and University Hospital Hradec Králové
MUDr. Mohammad Malekirad – Department of Psychiatry, Charles University, School of Medicine in Hradec Králové
Prof. MUDr. Klára Látalová, Ph.D. – Department of Psychiatry, Palacký University, School of Medicine in Olomouc and University Hospital Olomouc

CONTENTS

INTRODUCTION

A mental disorder usually occurs in one third of the population during the course of life. Mental disorders are mostly chronic and frequently life-long. On the whole, psychiatric disorders do not usually result in the patient's death, but may significantly deteriorate their quality of life. Their negative impact on the economy is also not negligible. Regardless of medical breakthroughs, contemporary treatment of mental disorders is far from satisfactory. We are able to treat; however, a complete and successful cure is still rare. One of the reasons is that we do not fully recognize the exact causes of mental disorders. That is why the treatment is normally symptomatic, rather than causal. Based on the ever-growing knowledge on the etiology of mental disorders, it will be possible to significantly improve therapeutic and preventive measures in the future, and thus mitigate the patients' suffering.

The aim of this textbook is to acquaint the readers with the most recent findings related to complex causes of mental illnesses. While preparing the text, we were amazed at how similar the etiology of individual mental disorders generally is to each other, regardless of strikingly different clinical presentations, e.g. dementia in contrast to schizophrenia or even anxiety disorders. In every case, some genetic background is challenged by specific environmental factors, and this gene-environment (GxE) interplay is represented by epigenetic changes. Individual mental disorders are only different from each other in the extent and specificity of particular etiological factors.

We wish you pleasant reading,

The authors

BASIC FACTORS RELATED TO ETIOLOGY OF MENTAL DISORDERS

The aim of this textbook is not and cannot be to explain in detail genetics, the molecular biology of the brain or statistics beyond the scope of clinical psychiatry. Nevertheless, in order to facilitate understanding of the text, we present a brief explanation of related basic professional and scientific terminology.

SINGLE-NUCLEOTIDE POLYMORPHISMS

A single-nucleotide polymorphism (SNP; pronounced "snip") is a difference in base pair that affects a single base pair. This term is used when referring to a variation that has a population frequency of 1% or greater. Single-nucleotide polymorphisms may be positioned within coding sequences of genes, non-coding regions of genes, or in the intergenic regions. The term "variant" with a qualifier about pathogenicity is preferred in clinical diagnostic testing over "mutation", although use is inconsistent in literature. However, strictly speaking in human genetics a mutation is a genetic variant of low population frequency in contrast to a polymorphism with an allele frequency of 1% or more. In contrast to mutations, which occur less frequently and generally have a negative impact on protein function, SNPs are not clearly detrimental.

In etiology of mental disorders, both common SNPs with small effect (a minor allele frequency of 0.05 or greater; a typical odds ratio of 1.05–1.15) as well as rare SNPs with large effect (a minor allele frequency 0.01–0.05) play a significant role.

Coding variant: A genetic variation in the protein encoding region of a gene. Coding variants are divided into non-synonymous and synonymous types, depending on whether the amino acid composition of the resulting protein is affected or not, respectively. Non-synonymous variants are further divided into missense and nonsense variants. In a missense variant, the amino acid sequence is altered but the protein is functional. However, in a nonsense variant the amino acid sequence is changed by the introduction of a premature stop codon and the resulting protein is non-functional. In a synonymous variant, the amino acid sequence of the protein is not affected.

Noncoding variant: A genetic variation that is not in the protein encoding region of a gene. However, these variants can be functional if they reside in and disrupt functional elements, for instance noncoding RNA sequences or regulatory sites (e.g. promoters, enhancers, suppressors, or splice-sites).

The term SNP is commonly used in research such as GWAS studies; see genome-wide association studies (GWAS) below.

COPY NUMBER VARIATIONS

Copy number variation (CNV) is the most prevalent chromosomal structural variation, a polymorphism, in which the number of copies of a relatively large chromosomal or DNA segment (> 1000 base pairs, usually can be measured from thousands to millions of bases) is repeated (duplicated) or deleted and varies between individuals. This results in an excessive production or non-production of the associated proteins, which may affect e.g. brain development or neuroplasticity. Copy number variations are associated for example with schizophrenia, intellectual disability, autism and epilepsy.

PLEIOTROPY

Pleiotropy occurs when one genetic variant influences multiple (two or more) phenotypic traits or has multiple biological effects. In other words, pleiotropy is the association of variant(s) in a single gene with several phenotypic effects.

EPISTASIS

It describes the interaction of variants at two or more genetic loci to produce phenotypes that are different from each variant's own impact. "Gene-gene interaction" or "genetic modifier effect" are both terms which are used to describe this process.

GENETIC ASSOCIATION STUDY

Traditional epidemiologic association studies and genetic association studies are comparable in certain aspects. Here, instead of searching for association between typical risk variable (e.g. child abuse) and disease outcome (e.g. development of anxiety disorders), a genetic association study looks for an association between genetic variable (i.e. genotype) and a specified condition (i.e. phenotype). The standard study design for association studies in general is case-control or nested case-control, where healthy controls are selected from the general population.

Candidate gene approach: Gene association studies in the past utilized a "candidate" gene approach, in which a genetic variant of interest was chosen based on the known or presumed biology of a disease, based on the association found in previous studies to a disease. In this strategy, one or a limited number of known variants are genotyped in a number of cases and controls, usually using PCR methods. Genetic association studies (i.e. the candidate gene approach) are already considered to be outdated and replaced by genome-wide association studies (GWAS).

The genome-wide approach: looks for an association between millions of variants and the condition of interest.

GENOME-WIDE ASSOCIATION STUDY

A genome-wide association study (GWAS) is an observational study approach which looks for and tests an association between hundreds of thousands of gene variants simultaneously, millions on a genome-wide set, and the condition of interest (i.e. a mental disorder). These are typically SNPs (>500 000 base pairs by convention) in different individuals to see if any variant is associated with a disease.

In polygenic and complex disorders, like most mental disorders, large numbers (dozens of thousands) of patients (i.e. cases) with the same number of healthy volunteers (i.e. controls) have to be studied. Such numbers are rarely obtained in single-center studies. That is why large international consortia for the genetic study of complex disorders were formed. Because a large number of statistical comparisons needs to be done, a strict limit is determined to establish a statistically significant result ($p < 5 \cdot 10^{-8}$). The scientific rationale for GWAS is the "common disease-common variant (CDCV)" hypothesis. It assumes that common diseases (e.g. schizophrenia or dementia) are caused by genetic variants in common genes (e.g. genes related to neurotransmitter activity), and each of these polymorphisms on its own contributes to the etiology in a very small extent. What decides is the total number of affected genes. The problem in GWAS is that even though a number of susceptibility variants (polymorphisms) were identified, they only explained a small proportion of the heritable risk (the "missing heritability" problem).

Complex diseases require novel genetic strategies to capture the "missing heritability", e.g. whole-exome sequencing or whole-genome sequencing. Several theories have been proposed to explain the "missing heritability" problem in GWASs, for example:

- Rare genetic variants of large effect, that will only be detected by whole genome sequencing
- Some of the variation may be explained by copy number variants (CNVs)
- Epistasis (gene-gene interactions), gene-environment interactions and epigenetics are also considered and studied.

GWASs have been used in psychiatry since 2007.

POLYGENIC RISK SCORE

The polygenic risk (PGR) score analysis is a statistical method which is used to summarize genetic effects among a group of SNPs which individually have only very small significant associations with a given disease. The polygenic risk score is constructed as a sum of scores of individual alleles weighted by their effect sizes. The polygenic risk score can be used as a tool to predict an individual's genetic risk of developing certain complex disorders, like mental disorders (for example schizophrenia) or complex polygenic traits.

However, the PGR scores may be "regionally biased" since they are commonly calculated based on data which are region specific due to the GWAS studies frequently comparing the common variants in the population of a given region, for example Europe. That is why it is important to increase the sample size of the population to achieve more reliable results.

PATHWAY ANALYSIS

Pathway analysis is a compelling method for data analysis in genomics. This is the analysis of the genes which may be associated with each other in some way in a biological pathway. For example, all the genes that are involved in or influence serotonergic neurotransmission. These genes are considered by the researchers as one unit in pathway analysis.

Pathway analysis is mostly used in gene expression analysis, however it is also applied to analyze variants, such as SNP data and it has been shown to be promising, for example, in the analyses of data from GWAS research. Here, it lets us to interpret variants in terms of the biological processes that the affected genes and proteins are involved in.

MICROBIOME

Microbiome is defined as the assembly of microorganisms and their genomes living in and on a human body. The human body is populated by trillions of symbiotic microorganisms – bacterial, viral, fungal, archaeal, parasitic, helminthic, and yeast. Microbiota influence the host organism via genetic, metabolic or other biological mechanisms. This host-microbiome influence is dynamic and bi-directional. Microbiome changes over time, and may be different in healthy (homeostasis) versus ill (dysbiosis) people. A given microbiome may be a cause as well as a result of a human disease, including mental diseases. The gut microbiota, composed of a complex community of bacteria (*Proteobacteria*, *Firmicutes*, *Actinobacteria*, and *Bacteroidetes*), archaea and fungi, is connected to the gut-brain axis via several pathways and mechanisms, for example via hypothalamic-pituitary-adrenal axis, toll-like receptors, brain-derived neurotrophic factor, tryptophan-kynurenine pathway, vagal pathway or epigenetic factors.

The abnormal activation of the HPA axis triggered by microbiota dysbiosis may result in disruption of the immune system leading to increased permeability of the gut barrier, release of pro-inflammatory cytokines and translocation of bacterial peptidoglycan into the brain. Toll-like receptors are elements that initiate inflammation in innate immune response. They are responsible for recognizing microbial motifs of both commensal and pathogenic microorganisms, especially in the gastrointestinal tract. Injured intestinal epithelium disrupts the activation of toll-like receptors resulting in gastrointestinal inflammation, which could lead to the development of neurodegenerative and other psychiatric disorders. As of epigenetic factors, gut microbiome-derived proteins and metabolites can directly interact with the epigenome or serve as cofactors for epigenetic enzymes that regulate DNA methylation and histone modifications. Dysbiosis (disruption to the gut microbiota homeostasis) can be caused by environmental factors such as dietary habits, urbanicity, pollution, stress or lifestyle.

ENVIRONMENTAL FACTORS

There are many different kinds of environmental factors involved in the etiology of mental disorders such as physical, biochemical, biological, nutritional, pharmacological, psychological and/or social ones. For example – DNA may be damaged by ionizing radiation or air pollution, cannabis or methamphetamine can induce schizophrenia in genetically pre-disposed individuals, corticosteroid medications can start or induce depression, abuse in childhood

increases the probability of anxiety disorders in adulthood, or an enormous social pressure to be busy, happy and successful triggers mental bulimia. Individual environmental factors will be described in the following text related to psychiatric diagnoses. In the gene-environment interplay, we generally distinguish between two possibilities:

Gene-environment correlation (rGE): a genetic influence on exposition to a given environment, e.g. a boy can decide whether he would go and play football or rather smoke marihuana. This decision is partially based on his genetic background.

Gene-environment interaction (GxE): a genetic influence on vulnerability to a given environmental factor, e.g. after an assault during a war, one soldier bravely continues to fight, while another one develops post-traumatic stress disorder.

EPIGENETICS

Epigenetics involves heritable or de-novo molecular changes that affect gene activity and expression, without alteration in the primary DNA sequence. Epigenetic effects may result from environmental factors, or be a part of physiological development. Epigenetics may affect both germinal or non-germinal cells. So far, four leading epigenetic mechanisms have been detected:

DNA methylation: In DNA methylation, methyl groups are added to the DNA molecule at CpG dinucleotides (cytosine-phosphate-guanine dinucleotides), termed CpG islands. When located in a gene promoter, DNA methylation mostly, but not always represses gene transcription. DNA methylation plays a crucial role in several key processes, such as genomic imprinting and X-chromosome inactivation. Methylation activity is maintained by DNA methyltransferase. DNA CpG hydroxymethylation, formylation and carboxylation or DNA non-CpG methylation have also been observed, their effect is not clear. Intragenic DNA methylation is supposedly modulating alternative splicing.

Histone modification: Various histone modifications control spacing of nucleosomes (basic structural units of DNA) and the degree to which they are condensed, which determines gene activity. Chromatin (DNA + histone proteins) exists in an inactivated, condensed state (heterochromatin) which does not allow transcription of genes, and in an activated, open state (euchromatin) which allows individual genes to be transcribed. This is controlled by biochemical histone modifications. Nearly two meters of DNA are packaged into the cell nucleus of the vast majority of eukaryotic cells. Histone proteins act to package DNA, which wraps around the eight histones, into chromosomes. A histone modification is a post-translational modification to histone proteins which includes methylation, phosphorylation, acetylation, ubiquitylation, and SUMOylation. These modifications impact on gene expression by altering chromatin structure or recruiting histone modifiers. The modifying enzymes involved in histone acetylation are called histone acetyltransferases and histone deacetylases. Histone acetylation increases gene expression in the relevant parts of DNA. The role of other histone modifications is less known.

Non-coding RNA: These are regulatory RNAs such as micro RNAs and long non-coding RNAs. MicroRNAs (miRNAs) are small, endogenous, noncoding RNAs, approximately 18–24 nucleotides in length which regulate gene expression. Micro RNAs affect the protein levels via silencing of the relevant target mRNAs by binding to it and inducing mRNA degradation or inhibiting translation, so no protein is assembled and hence gene expression is