Molecular epidemiology; New but impressive

Maryam Honardoost[1]*, Azam Rajabpour[2], Ladan Vakil[2]

Précis

Molecular epidemiology is a subdivision of medical science and epidemiology that emphases on the involvement of potential environmental and genetic risk factors, recognized at the molecular level, to the etiology and avoidance of sickness through populations. This arena has developed from the combination of molecular biology and traditional epidemiological research. Molecular epidemiology can improve our knowledge about the precise pathogenesis of disease through recognizing particular pathways that affect the risk of developing the disease. Furthermore, it tries to find how the collaborations between genetic characteristics and environmental exposures works in disease occurrence.

Introduction

Molecular epidemiology is a division of medicine that integrates innovative laboratory methods into epidemiology to recognize the etiology of disease and facilitate intervention. It is progressively utilized to elucidate interactions between genetic, environmental and other factors, and to identify susceptible populations and individuals. This course will cover conceptual and methodological issues in molecular epidemiology including the application of biomarkers to study disease causation, risk assessment, and prevention.

The science of Epidemiology entails applying classic epidemiological methods to show the distribution of disease in the population and the art side of that is interpretation of findings (1).

Molecular epidemiology emphasizes the role of each genetics and environmental factors that influence disease process at molecular levels. The outcome of this work is the elucidation of disease etiology, distribution pattern and penetrance in families and population (2).

This field is composed of molecular biology and basic epidemiology and with detection and introduction of the molecular pathway and specific genes that play role in disease risk, help us to understand disease pathogenesis.

Molecular epidemiology is a new opportunity for the epidemiologist and other medical scientist to understand accurate disease mechanisms and make public health recommendations for disease prevention and treatment (1).

History

The phrase "Molecular epidemiology" for the first time was introduced by Kilbeurne in 1973 in the article "Molecular Epidemiology of Influenza" (2). The word became more formalized with the preparation of the first book on Molecular Epidemiology: "principles and practice" by Schulte and Perera. This book is about impact of advances in Molecular research that have given use age of its and enable the measurement and explanation of biomarkers as vital tools to connect traditional and epidemiological research strategic,and also to understand fundamental molecular mechanisms of disease in the population.

Science kilbourne's application of the term, "Molecular Epidemiology", has led to steady growth in the use of the word in about 17000 papers have been published till now. However this numbers would not ordinarily deemed by a large number, nor this include the vast explosion of scientific literature on biomarkers, genetics enzymology as well

_____

Corresponding author: Dr Maryam Honardoost, honardoost.m@iums.ac.ir

[1.] Endocrine Research Center, Institute of Endocrinology and Metabolism, Iran University of Medical Sciences, Tehran, Iran.
[2.] Molecular Medicine Department, Faculty of Medical Science, Qazvin Medical University, Qazvin, Iran.

↑What is "already known" in this topic:
Molecular epidemiology increases our knowledge about the pathogenesis of disease by identifying vital molecules, genes and pathways that have effect on the risk of disease development.

→What this article adds:
A brief introduction to molecular epidemiology as a new opportunity for medical scientist to understand accurate molecular base of disease and make public health recommendations for disease prevention and treatment.

as molecular and cell signaling of disease, all of which lend themselves, but which may not be recognized as to the ideas and philosophies of molecular epidemiology.

Molecular Epidemiology: interrelationship ring

For an epidemiologist, molecular epidemiology is included of each epidemiologic study that utilizes special tools to study some biologic factors such protein profile, metabolites or, even new genes in some patients. This science facilitates disease etiologies detection, (especially chronic diseases), the introduction of disease determinant, prevention of disease incidence and improvement of health conditions at public level. Molecular epidemiology helps the acquisition of information about the disease and in such studies, with biomarker evaluation and genetic tools for hereditary and acquired factors, ambiguous cases would be clarified (1-3).

Molecular epidemiology mainly focuses on genetic and environmental risk factors contribution and the interaction between them at molecular or cellular stages. Complete human DNA sequencing as the main product of Human Genome Project (completed in 2003) was a breakthrough in this field. Moreover, recent advances in genotyping and development of high-throughput technologies which evaluate DNA, RNA, Protein or metabolic components of a person, have provided the foundation to more comprehensively examine most of possible biologic variations related to disorder risk factors and disease pathway (Fig. 1). High-throughput technologies have, nevertheless, extended our understanding of disease phenotype-genotype association and helped to look for biomarkers of disease and use them to recognize the fundamental mechanisms of disease in populations. This approach will help the clinical practitioner to customize treatment in individual basis. For example, currently, many laboratories offer NGS panel for many disorders to accelerate early detection of syndromes and to detect

pharmacogenetics marker for choosing better treatment strategy. Therefore, Molecular Epidemiology is a multipurpose science that makes novel projections possible.

Bias in molecular epidemiology studies

Random error as a major problem in such studies, drown from an inadequate number of subjects involved in each research. The core reason for a small research lies with financial and logistical constraints. Linkage disequilibrium, population stratification, pleiotropy and canalization and gene-environment interactions are other important source of bias in molecular epidemiology studies that should be avoided or reduced by using suitable study design. Sensitive and specific methods can help to overcome this problem to some extent (4).

Sensitive part of these molecular studies is the selection and properly designed study and identification of capabilities as well as limitation of a designed study. All the designed studies potentially have bias sources, and therefore, it is important to identify the source and minimize them to some extent. The main aim of molecular epidemiology is that to determine target population and interpret them exactly. The interpretation should arise from the studies on the same population (1, 2). The inaccurate relationship between parts of a study can influence on its interpretation, especially if data collected randomly, even in cases such as genetic or molecular factors. This incorrect relationship would eventuate different results.

Hence, for sample collection and accurate information, awareness of study population, seems to be essential. Like other epidemiologic studies, main bias could affect biomarker-based research. Selection bias rises from lack of study group comparability (e.g. cases and controls). For instance, the exposed case might be more or less probable to participate in a study then exposed controls or volunteers may consitute people with a family history of
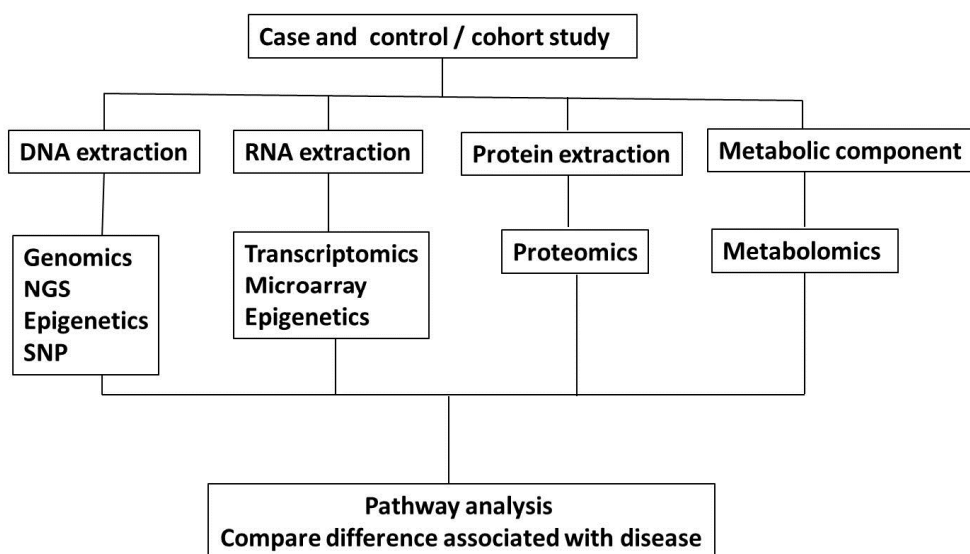

Fig. 1. Methods used to study biological variations in molecular epidemiology. Abbreviations: NGS, next generation sequencing; SNP, single nucleotide polymorphism.

tumors (1, 5).

Information bias contains misclassification of applicants with respect to disease or exposure studies. In biomarker-based studies, information bias covers the issue of validity, reproducibility, and stability of markers (5).

Confoundings are other types of bias sources. One characteristic of molecular epidemiology investigations is the reasonably large number of variables on exposures, disease and effect modifiers. By doing so, the chance of statistically significant results generated by chance, can increase dramatically. Furthermore, a matching method based on the score of Propensity can be used to assure the same distribution of confounders in case and control group.

The main source of bias in studies about genetic factors like metabolic polymorphisms is due to the lack of attention to the selection of participants in each project. Because it is impossible to choose a participant related to the favorite genetic factor. For instance, the correlation between CYP2D6 gene and lung cancer, in which it is rejected by some study due to inappropriate selection of control group (3).

Moreover, the variability of reports for biomarker levels is due to intergroup variability (case versus controls). And this event happens in molecular epidemiology because of misclassification and individual differences caused by genetic and environmental factors.

Statistical Power of molecular epidemiologic studies

Power calculations are vital for improving study design, because investigators have no control over the real genetic makeup that causes a phenotype, so, they try to control many phases of study design, such as the sample selection, phenotype measurement, and the methods of variants analysis, and the statistical method to be used. It is valuable to enhance the statistical power of a study, given the restrictions enforced by nature or by limitations in possessions. Many factors influence the statistical power of molecular studies, but only a few are under the researcher's control. Manipulating factors such as selection of study subjects, sample size, methods of measurements, and data quality control or using suitable statistical software, such as SimRare82 or SEQPower, make the investigators increase the statistical power of their study within the limitations of available resources. Also accurate data quality control can filter out challenging variations, helps to decrease false-positive relations and to increase power. Statistical analyses having enough account of covariates and potential confounding factors also help to improve statistical power (4).

Biomarkers: Determinant of Molecular Epidemiology

Application of biologic approaches, like biomarkers, to evaluate specific variables in epidemiologic studies is not a new approach. For decades, biomarkers have been used commonly in epidemiologic studies of infectious and cardiovascular diseases (3). For instance, increased levels of cholesterol, as a biomarker for cardiovascular disease and ischemia, have been proven in Molecular Epidemiology.

The recent development of biomarker-based studies has altered epidemiologic studies process in chronic disease especially cancer epidemiology. Consideration of molecular epidemiology as an alternative for epidemiologic studies seems to be essential because its main aim is identification of disease determinants (risk factor or protective factor). Information bias along with other biases seen in case-control or cohort studies are reduced in such research.

Molecular epidemiology studies are subjected to a specific pattern. In other words, they include the studies which measure biomarker and risk factor and their resulted confounders relations and modifiers (3, 5).

Study analysis, interpretation of molecular epidemiology studies, is similar to epidemiologic study biomarkers distinguished as disease-causing markers or susceptibility and hence they are categorized in their groups of exposure, disease, and susceptibility. Exposure group is consisting of proteins or DNA adducts even viral DNA and urinary metabolite, and disease group includes gene mutation, chromosomal aberration, and neoplastic condition. Susceptibility group is including metabolic polymorphism and DNA repair activity.

However somewhat arbitrary is probable between these cases, such as, chromosomal aberrations which have been used to observe the exposure to environmental carcinogens for decades, so, they can be categorized as exposure biomarkers. However, current evidence shows a role of chromosomal aberration in predicting disease risk.

Recently epigenetic factor such micro RNA (miRNA) research has expanded into a huge number of disease areas, it has become obvious that expression levels of specific miRNAs are altered in numerous diseases (6). By itself, the function of these small non-coding RNAs as biomarkers have become noticeable and exploiting and has become a focus for researchers around the globe.

Obvious point of biomarker evaluation is the reliable sensitivity and specificity of the methods which determine disease determinants. Also, authentic and standard methods of sample collection would be helpful in correct biomarker reports.

Molecular Epidemiology and cancer

Molecular epidemiology of cancers refers to an attitude in which innovative laboratory methods are used with analytical epidemiology to find, at the molecular level, specific exogenous causes and/or host factors that have a role in human cancer causality. Molecular epidemiology is useful to determine genetic variants or polymorphism related to cancer (as disease or exposure determinants). A considerable point of these genetic variants is independently related to regimen, environmental conditions which can alter the risk of disease and make their studies ambiguous. Despite the fact that these studies are limited in the study of acquired changes, but a better understanding of potential risk factors is available.

Two aspects of dosimetry (estimation of the actual amount of biomarkers) and susceptibility studies (genetic

variants) are known as the independent predictors of sickness or relationship modifiers of exposure and cancer.

Molecular epidemiology of cancer studies provide the potential for clarifying the molecular level of carcinogens... nonetheless, identification of prone subsets of the population based on polymorphisms in genes related to cancers, has the potential to describe more noticeably those factors that might increase cancer risk among some people.

Advances in human genomics determines susceptible subgroups bared based on genetic variant in various pathways such as carcinogens metabolism, genome repair, cell cycle and immune status that increases our knowledge about cancer etiology and with development of statistical analysis, the best strategies for prevention would be available (7).

Examination of the relationship between polymorphism of CYP1A1 gene and tobacco-related cancers in the upper aerodigestive tract or relationship between a special variant of CYP and increases the risk of breast and head and neck cancer especially in people with inherited glutathione-s-transferase (GST) M1 (9-10). Also, other approach is XRCC1 variant alleles that display inferior base excision repair capacities that can be related to greater risk of oral and pharyngeal cancer (11). Moreover, the interrelationship between genes which regulate cell growth such as H-ras, P53, HLA genes, immune system, cytokines, and cancer is the aim of these studies (12-14). However, the majority of these studies in the field of cancer problems is the surveys based on previously studied genes and new genes are not these studies subjects. It is better that these studies work on new functional genes which involved in the favorite pathway.

Molecular epidemiology and infectious disease

Molecular epidemiology is science which molecular biology helps in identification of distribution patterns of disease in the population. Pathogen genome project has helped multilocus sequence typing of pathogen isolates and ecological fitness survey and virulence patterns in disease. A better understanding of a pathogen distribution outside of its host and finding new disease-causing pathogen would be possible by such these researchers (15, 16). For example, a molecular epidemiologic study has provided the genomic evidence of HIV-1 predominance genotype infection associated with injection drug addicts in Iran (17) or Shah-Hosseini, et al investigated the possible association of KIR-HLA genotypes with HBV infection and base on their result, they suggested a potential influence of the immune cells' activating phenotype that clear viral infection in Iranian HBV positive patients (18).

By new methods like pyrosequencing and molecular marker detection, we are able to translate complex pathways of disease and find the distribution routs.

Another helpful work is positional cloning of one infections disease susceptibility genes such as Slc11a1 or Nranp1 in the mouse to genome-wide scans in human multicase families and the identification of potential disease genes are all helpful ways in molecular epidemiology

studies infections disease revolution. The phylogenic studies between resent isolates and ancestors with phylogenic trees are effective tools in study of pathogen and alteration level of them (15).

In all, comparative sequence analysis of pathogen strains and functional genomic studies are in progress, providing new vision into disease susceptibility.

Consideration for future

Despite the fact that about 30 years passed since the term Molecular epidemiology was exposed and the importance of molecular methods in epidemiologic research is approved, but there are some limitation. Inconsistency and inconclusiveness of the result in many molecular epidemiology studies are due to biases and errors flew in study design. Small size results in insufficient power to detect association and is a usual problem in molecular epidemiology reserches , because of the interaction between genotype and exposure. The fact that, genotype alone may not represent enzyme activities and the importance of polymorphism on disease incidence in the disease process is still one of the most problems of these surveys. The function of many polymorphisms is not cleared and evaluation of such genes in cancer is questionable. Procedures that provide a real-time evaluation of tens of hundreds of genetic variance in a large number of the sample will be valuable in molecular epidemiology studies of cancer. With the purpose of designing effective cancer prevention programs, we must recognize the main characteristics of the disease.

Molecular epidemiology surveys are designed to help researchers to conduct etiologic study and increase our understanding about the disease determinants. Moreover, molecular epidemiology will elucidate the mechanism of cancer formation and progression while also finding a susceptible population can be targeted for cancer prevention. Several factors which affect the gene-environment interaction, can be changeable and might be a suitable point for primary prevention. The molecular epidemiology covers principles in the biomarker selection and validation, research design and statistical analysis methods for gene-environment interactions, sample collection, storage, and banking, and recent laboratory methods for biomarker analysis.

Conflict of Interests

The authors declare that they have no competing interests.

References

1. Zhebrun AB. From molecular to genomic and meta-genomic epidemiology. Zh Mikrobiol Epidemiol Immunobiol. 2014 May-Jun;(3):91-100
2. What is Molecular Epidemiology? (http:///www.pitt.edu/~kkr/molepi.html). Molecular Epidemiology Homepage. University of Pittsburgh. 28 July 1998. Retrieved 15 January 2010.
3. Malaguarnera R, Vella V , Pellegriti G, Belfiore A. Editorial: Clinical and Molecular Epidemiology of Thyroid Cancer of Follicular Origin.Front Endocrinol (Lausanne). 2018 Mar 12;9:67.
4. Sham PC, Purcell SM. Statistical power and significance testing in large-scale genetic studies. Nat Rev Genet. 2014

May; 15(5):335-46.

5. Schulte Paul A, Perera Frederica P. Molecular Epidemiology: Principles and Practice. (1993) Academic Press. pp. 588.

6. Honardoost M, Sarookhani MR, Arefian E, Soleimani M. Insulin resistance associated genes and miRNAs. Appl Biochem Biotechnol. 2014 Sep; 174(1):63-80.

7. Vineis P, Perera F. Cancer Epidemiol Biomarkers Prev. Molecular epidemiology and biomarkers in etiologic cancer research: the new in light of the old. Cancer Epidemiol Biomarkers Prev. 2007 Oct;16(10):1954-65.

8. Ezzeldin N, El-Lebedy D, Darwish A, El Bastawisy A, Hassan M, Abd Sh, et al. Genetic polymorphisms of human cytochrome P450 CYP1A1 in an Egyptian population and tobacco-induced lung cancer. Gene Enviro. 201739:7.

9. Yi-Wen Z, Mei-Hua B, Xiao-Ya L, Yu C, Jing Y, Hong-Hao Z. Effects of Oridonin on Hepatic Cytochrome P450 Expression and Activities in PXR-Humanized Mice. Biol Pharm Bull. 2018;41(5):707-712.

10. Karahalil B, Bohr VA, Wilson DM. III Impact of DNA polymorphisms in key DNA base excision repair proteins on cancer risk. Hum Exp Toxicol. 2012 Oct;31(10):981–1005.

11. Reyes-Gibby CC, Wu X, Spitz M, Kurzrock R, Fisch M, Bruera E, et al. Molecular epidemiology, cancer-related symptoms, and cytokines pathway. Lancet Oncol. 2008 Aug; 9(8): 777–785.

12. Ahmad D. Epidemiology Study on P53 (Rs1614984) C>T Mutation in Cigarette Smokers. Braz Arch Biol Technol. 2017;60.

13. Mi Kim E, Jung CH, Kim J, Hwang SG, Park JK, Hong-Duck Um. The p53/p21 Complex Regulates Cancer Cell Invasion and Apoptosis by Targeting Bcl-2 Family Proteins. Cancer Res. 2017;77(11); 1–9.

14. Snow M. The contribution of molecular epidemiology to the understanding and control of viral diseases of salmonid aquaculture. Snow Veterin Res. 2011;42:56

15. Eybpoosh S, Haghdoost AA, Mostafavi E, Bahrampour A, Azadmanesh K, Zolala F. Molecular epidemiology of infectious diseases. Electron Physician. 2017 Aug; 9(8):5149–5158.

16. Shah-Hosseini A, Jafari M, Mohammadi A, Sanaei R, Alavian SM, Doosti-Irani A, et al. The impact of KIR-HLA genotype on hepatitis B virus clearance in Iranian infected individuals. Med Microbiol Immunol. 2017 Oct 14.

17. Jahanbakhsh F, Ibe S, Hattori J, Monavari SH, Matsuda M, Maejima M, et al. Molecular epidemiology of HIV type 1 infection in Iran: genomic evidence of CRF35_AD predominance and CRF01_AE infection among individuals associated with injection drug use. AIDS Res Hum Retroviruses. 2013 Jan; 29(1):198-203