Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
INTRODUCTION

Population sciences such as epidemiology focus on studying whole populations rather than individuals. Through studying environmental, genomic, and social factors that affect human health, population level interventions can be identified. The social sciences focus on studying society and human behavior through fields such as anthropology, economics, law, psychology, and sociology. The study of the ethical, legal, and social implications (ELSI) of genomics plays an important role in applying genomics to population health.

Although the fields of population genomics and genomic medicine look to prevent or treat disease through different perspectives, they can act complementarily to enhance overall health outcomes for both individuals and populations at large. Population genomics seeks to integrate knowledge from genomic, population, and social sciences to improve population health.

The genomic sciences focus on studying whole genomes, such as the entire DNA sequence making up the human genome. Through studying genomics, genetic variants influencing human health can be identified. Studies of particular genes can then further elucidate the function of genetic variants. The three major disciplines contributing to population genomics (genomics, population science, and social sciences) explore two cross-cutting issues: global health, and population versus individual health. This aspect of the emerging field of population genomics is discussed in this chapter using specific examples from diverse diseases such as breast cancer, colorectal cancer, bronchial asthma, Crohn disease, Alzheimer dementia, and cystic fibrosis.

BREAST CANCER

Around one in nine women worldwide develop breast cancer during their lifetime. A family history of breast cancer along with ovarian cancer may be encountered in approximately 5% of the affected women. A number of disease-causing mutations in the two major genes, BRCA1 and BRCA2, are now recorded worldwide. In addition, several variants may occur in 5–10% of breast cancer cases; these variants are found in less than 1% of the general population [1]. Specific populations, such as those of Ashkenazi Jewish descent, have an increased incidence of BRCA1 or BRCA2 variants. Two variants in BRCA1 and one in BRCA2 are found at a rate 5 times higher in Ashkenazi Jews than in the general population [2,3].

For these reasons, family members of those with known BRCA1/2 variants or those with a family history of breast cancer may be offered genetic testing. Men with BRCA1/2 variants are also at an increased risk of developing breast cancer [4,5]. Thousands of variants have been discovered in BRCA1 and BRCA2, yet only a minority have a known deleterious effect [6]. Genetic testing therefore has the possibility of finding a variant of unknown effect, for which the functional significance is unclear. One recent study found that 10% of women undergoing BRCA1 and BRCA2 testing receive an ambiguous test result because of the detection of a variant of unknown significance [7]. Thus although deleterious variants are known to increase the risk of developing breast cancer approximately fivefold, deciding how to react to variants of unknown effect can be challenging for all involved including both clinicians and patients [8].
COLORECTAL CANCER

Genomic information can be used clinically to inform disease risk, diagnosis, drug selection, and drug dosing. Colorectal cancer provides a good example of an area where population level screening along with genomic medicine approaches are coming together to improve overall population health.

Over 1 million individuals are diagnosed with colorectal cancer each year worldwide, accounting for approximately 9–10% of cancer diagnoses in 2008 [9]. It is the third leading cause of cancer-related death in the United States and the fourth worldwide [9,10]. In colorectal cancer, a patient’s genomic information can be used to determine risk of inherited colorectal cancer syndromes, whether certain biological agents will work in specific patients, and what starting dose to use on specific chemotherapeutics.

As many as 20–25% of colorectal cancer cases have a family history of colorectal cancer (two or more first-degree relatives with colorectal cancer), yet only 5–6% have an established familial genetic syndrome with a known genetic variant [11,12]. Of those with established familial genetic syndromes, approximately 3% will be diagnosed with Lynch syndrome (including variants in the genes MLH1, MSH2, MSH6, PMS2, and EPCAM) and 1% with familial adenomatous polyposis (including variants in APC and MUTYH) [11]. Individuals with a family history of colorectal cancer have a two- to threefold greater risk of developing colorectal cancer than the general population and thus genetic testing for an individual with a known family history has substantial public health benefit [11]. Those with a known family history of colorectal cancer are also recommended for screening at younger ages, typically 10 years younger than the onset of the youngest case in their family.

Genetic testing is also used to determine treatment options in colorectal cancer [11,12] (Table 23.1). For example, genetic variants that make KRAS constitutively active have been shown to provide resistance to monoclonal antibodies directed against the upstream epidermal growth factor receptor (EGFR), because both are components of a cellular pathway leading to abnormal cell growth and cancer. Thus cetuximab and panitumumab (anti-EGFR antibodies) are given only to individuals with normally functioning KRAS, where blocking EGFR can have an effect [11].

Pharmacogenomics can also be useful in determining drug dosage for colorectal cancer. For example, the FDA recommends testing for UGT1A1 variants when administering irinotecan, because individuals homozygous or heterozygous for the UGT1A1*28 allele are at increased risk of developing neutropenia and severe infections [13]. Individuals with inactivating UGT1A1 variants are therefore recommended to be started at a lower dosage of irinotecan to reduce the risk of neutropenia [13].

In addition to modifying drug dosing, pharmacogenomic information can also be used in drug selection to choose agents more likely to give a beneficial response based on a patient’s genetically driven ability to metabolize them. For example, in patients of Asian ancestry given carbamazepine (used to treat epilepsy and bipolar disorder) the HLA-B*1502 allele has been associated with Stevens-Johnson syndrome/toxic epidermal necrolysis, a life-threatening skin condition. This allele can be found in over 15% of the population in some regions in Asia including Hong Kong, Thailand, Malaysia, and parts of the Philippines, and is very rare in other populations outside Asia [14] (Table 23.2).

Within populations of Asian ancestry, there can also be great variation, such as is seen within China where the HLA-B*1502 allele prevalence varies from 0 to 36%, depending upon ethnicity. Thus the US Food and Drug

---

**TABLE 23.1** Pharmacogenomic Variants in Colorectal Cancer [11].

| Function                        | Consequences                                                                 |
|---------------------------------|-----------------------------------------------------------------------------|
| BRAF                            | Downstream pathways constitutively active                                   |
| ERCC-1                          | DNA excision repair                                                         |
| Interleukin 8                   | Increased VEGF expression                                                  |
| KRAS                            | Downstream pathways constitutively active                                   |
| Microsatellite instability       | Reduced DNA repair                                                         |
| TSER                            | Increased or decreased thymidylate synthase, depending on variant           |
| UTCA1                           | Responsible for metabolism of irinotecan                                   |
| VEGF                            | Increased VEGF expression                                                  |

**Function**

- **BRAF**: Downstream pathways constitutively active
- **ERCC-1**: DNA excision repair
- **Interleukin 8**: Increased VEGF expression
- **KRAS**: Downstream pathways constitutively active
- **Microsatellite instability**: Reduced DNA repair
- **TSER**: Increased or decreased thymidylate synthase, depending on variant
- **UTCA1**: Responsible for metabolism of irinotecan
- **VEGF**: Increased VEGF expression

**Consequences**

- Resistance to anti–epidermal growth factor receptor monoclonal antibodies
- Resistance to platinum-based chemotherapy drugs
- Increased cancer recurrence
- Resistance to anti–epidermal growth factor receptor monoclonal antibodies
- Improved prognosis
- Response to fluorouracil reduced/increased (negative relationship)
- Dosing for irinotecan
- Increased cancer recurrence
Administration (FDA) recommends genetic testing for the HLA-B*1502 variant before prescribing carbamazepine for patients of Asian ancestry and for those with one or two copies of the variant only treating with carbamazepine when the benefits outweigh the risks of the drug [14].

### Bronchial Asthma

Population variation is an important consideration when studying common complex conditions that are influenced by multiple genetic, environmental, and social risk factors, such as bronchial asthma. Over 300 million individuals of all ages have asthma worldwide [15]. Prevalence estimates can vary greatly by ethnicity; however, from 2% to 33% [16]. In the United States, prevalence ranges from approximately 8% in European Americans to 12% in African Americans and 7% in Hispanic Americans [17]. Within admixed populations, such as Hispanic Americans, even greater variation can be seen when populations are further stratified, with Mexican American populations around 6%, whereas Puerto Rican populations are closer to 19% [17]. Genetic studies have shown that at least some of this variation is caused by differences in genetic variants, with 35–80% of the variation in asthma heritability explained by genetic factors [18,19]. For example, variants in ADAM33 have been seen in European, African American, and some Hispanic populations, but not in other European American, Mexican, Puerto Rican, and Korean populations, all of which found different variants in ADAM33 associated with asthma [20] (Fig. 23.1).

Studying the interplay between environmental, genetic, and social risk factors is also critical to understanding the etiology of this complex disease. For example, the effect of air pollution on asthma case reports is modified by genetic factors as well, showing potential gene–environment interactions. A key measure of air pollution is PM10, the concentration in parts per million of particulate matter 10 μm in diameter or less, which can penetrate and irritate small airways. PM10 has been shown in multiple epidemiological studies to be an independent risk factor for increased respiratory symptoms including asthma [21–25]. Similarly variants in over 100 genes have been associated with asthma in genome-wide association studies (GWASs) [26,27,28]. Looking at the two risk factors together, however, reveals a potential gene–environment interaction where variants in GSTP1, SOD2, and NFE2L2, all related to oxidative stress pathways, were also associated with increased hospital admissions for asthma-related symptoms during days with high PM10 levels [21].

### CROHN DISEASE

Genomics can also be used to help identify and better define environmental risk factors in population studies. For example, genomic data profiling the bacteria inhabiting the human gut, or gut microbiome, has revealed differences in the bacterial populations present in individuals with Crohn disease (a form of inflammatory bowel disease) [28]. The genomic signatures of the gut microbiome in patients with Crohn disease shows some bacterial populations to be decreased, whereas others are more abundant [28–31].

### ALZHEIMER DEMENTIA

Many genetic loci have also been associated with multiple phenotypes, as evidenced in the National Human Genome Research Institute Catalog of Published GWASs [27,32]. (Fig. 23.2). Such pleiotropic genes (genes associated with multiple phenotypes) can present additional challenges when considering the ELSI of returning genetic testing

---

**TABLE 23.2 HLA-B*1502 Allele Prevalence in Worldwide Populations [46].**

| Population     | HLA-B*1502 Allele Prevalence |
|----------------|------------------------------|
| China          | 0–36%                        |
| Indonesia      | 11–17%                       |
| Malaysia       | 21–6%                        |
| Vietnam        | 14%                          |
| Thailand       | 8–9%                         |
| India          | 0–6%                         |
| Singapore      | 6%                           |
| Taiwan         | 4–6%                         |
| USA            | 0–4%                         |
| South Korea    | 0.2–2%                       |
| Australia      | 0–1%                         |
| Japan          | 0.10%                        |
| Germany        | 0%                           |
| Brazil         | 0%                           |
| Bulgaria       | 0%                           |
| Burkina Faso   | 0%                           |
| Cuba           | 0%                           |
| Ireland        | 0%                           |
| Italy          | 0%                           |
| Mexico         | 0%                           |
| Morocco        | 0%                           |
| Oman           | 0%                           |
| South Africa   | 0%                           |
FIGURE 23.1 Example of the variation in allele incidence by population for rs2280089 in ADAM33. The A allele has been previously associated with predisposition to asthma and bronchial hyperresponsiveness in populations from the United States, United Kingdom, and China [48–51].

FIGURE 23.2 The National Human Genome Research Institute genome-wide association study (GWAS) catalog showing that many genetic loci are associated with multiple phenotypes. APOE on 19q13.32 is highlighted along with examples of the disease phenotypes associated with the gene [27].
results. For example, variants in \textit{APOE} are associated with multiple phenotypes, including Alzheimer dementia, cholesterol level, coronary disease, C-reactive protein, hyperlipoproteinemia type III, low-density lipoprotein level, macular degeneration, and response to statin therapy [33]. The \textit{APOE*ε4} variant in particular has been associated with increased risk for developing both Alzheimer dementia and atherosclerosis along with a protective effect against developing macular degeneration [33].

**Cystic Fibrosis**

In many ways the family serves as an intermediary between individual- and population-level views of health. It is an important viewpoint that should be considered in population genomics, because genomic information is inherently relevant not only to the individual tested, but also to their family members with whom they share a large proportion of their genetic variants. How and with whom such family-related health information can or should be shared is an important consideration for advancing both individual and family health.

The availability of genomic information is also blurring the line between population and individual-level views of health. For example, genetic testing for cystic fibrosis spans population screening–based carrier, prenatal, and newborn tests to individualized genomic medicine–based diagnostic and pharmacogenomic testing for treatment selection. From the population screening perspective, genetic testing is offered to prospective parents of European decent and others who may be at increased risk of having a child affected by cystic fibrosis, because the prevalence of cystic fibrosis is highest in Northern Europe [34].

Over 1500 variants have been found in the \textit{CFTR} gene, but the functional significance of many is unknown, with the most common variant associated with cystic fibrosis being ΔF508 [34]. In 2012, the FDA approved ivacaftor, the first drug to treat a specific cystic fibrosis variant, G551D in \textit{CFTR} [35,36] (Table 23.3). The G551D variant impairs the ability of the \textit{CFTR} channel to open [34,35,37]. Ivacaftor functions by increasing the likelihood of the \textit{CFTR} channel being open, improving chloride transport and restoring the function of the \textit{CFTR} gene [35–37]. As the cost of genomic sequencing continues to drop and electronic health records improve, the cost of collecting and interpreting genomic data may fall below the cost of conducting individual genetic tests, further blurring the line between clinical and public health data.

**CROSS-CUTTING ISSUES OF POPULATION GENOMICS**

Although all three of the population sciences contributing to population genomics work together, there are also some issues that more broadly span the field of population genomics and its relationship to medicine and public health. Touched upon in many of the examples discussed, it is important to consider the broader implications of population genomics to global health and how population and individual-level views of health can work together to improve health worldwide.

Cardiovascular disease is a leading cause of death worldwide, with over 13.5 million deaths from ischemic heart disease, stroke, or another form of cerebrovascular disease in 2008, and is highly amenable to study using population genomics techniques [38]. For example, adding rs10757274 genotyping to the Framingham risk score improved its ability to determine individuals who would suffer later cardiovascular events independent of family history [39]. Such models can be used to screen populations to determine individuals at increased risk of disease and recommend further testing and individualized genomic medicine.

Whereas chronic conditions such as cardiovascular disease make up the majority of deaths in the developed world, infections are still a major health concern within developing countries and are equally amenable to study using population genomics. Genomics has made possible the rapid identification of the organisms causing recent pandemic outbreaks including H1N1 and severe acute respiratory syndrome, as well as identifying the source of foodborne illness. The availability of genomic sequence information on malaria parasites, mosquito vectors, and their human hosts are all being leveraged to produce more rapid diagnosis and better drugs, vaccines, and intervention strategies to fight malaria [40,41].

To maximize the benefit of population genomics advances to global health, it is also important to include multiple populations of diverse age, ethnicity, and gender in disease research. As evidenced by the example of asthma genomics, the prevalence of disease can be highly variable across ancestral groups and genetic variants often vary in incidence as well. Thus although a single pathway may be implicated in disease across many populations, the most common variant in each population may lie in different genes or gene regions.

Local environmental and social factors that impact disease and population health should also be incorporated into

| TABLE 23.3 National Heart, Lung, and Blood Institute Exome Sequencing Project Results for African-American and European-American Participants for the G551D Variant (rs75527207) Associated with Cystic Fibrosis [47]. |
|-----------------------------|-----------------------------|
| Allele Count A | Allele Count G |
| African-American | 0 | 4406 |
| European-American | 18 | 8582 |
studies of population genomics to produce the most complete picture of disease etiology. For example, the prevalence of type 2 diabetes mellitus is increasing globally and has been associated with multiple genetic (more than 60 genes to date), epigenetic (such as methylation or histone modification), environmental (such as diet), and social factors (such as exercise), all of which contribute to this complex disease [42]. The prevalence of type 2 diabetes mellitus varies by country from approximately 5–29%, with risk alleles such as the C allele in rs11196205 decreasing in incidence from sub-Saharan Africa to Asia (Fig. 23.3) [43,44]. Effects of other risk factors also vary across different populations, with the relative risk of type 2 diabetes mellitus for each 5 kg/m² increase in body mass index; for example, being 2.4 in Asian Americans, 2.2 in Hispanic Americans, 2.0 in European Americans, and 1.6 in African Americans [45].

**SUMMARY**

In this chapter we have explored how the integration of genomic, population, and social sciences in population genomics can improve health through examples in pharmacogenomics, population variation, and genetic pleiotropy. We have also investigated cross cutting issues in global health and population versus individual health where population genomics can play a crucial role in the translation of genomic health discoveries worldwide and population screening can work together with genomic medicine to provide the greatest health benefit to both individuals and populations at large. Thus, multidisciplinary research in population genomics, can improve clinical care through understanding of the genetic variation in populations that contributes to complex disease.

**REFERENCES**

[1] Schwartz GF, Hughes KS, Lynch HT, et al. Proceedings of the International Consensus Conference on Breast Cancer Risk, Genetics, & Risk Management, April, 2007. Cancer November 15, 2008;113(10):2627–37.
[2] Struwing JP, Hartge P, Wacholder S, et al. The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews. N Engl J Med May 15, 1997;336(20):1401–8.
[3] Warner E, Foulkes W, Goodwin P, et al. Prevalence and penetrance of BRCA1 and BRCA2 gene mutations in unselected Ashkenazi Jewish women with breast cancer. J Natl Cancer Inst July 21, 1999;91(14):1241–7.
[4] Thompson D, Easton DF. Cancer incidence in BRCA1 mutation carriers. J Natl Cancer Inst September 18, 2002;94(18):1358–65.
[5] Cancer risks in BRCA2 mutation carriers: the Breast Cancer Linkage Consortium. J Natl Cancer Inst August 4, 1999;91(15):1310–6.
[6] An open access on-line breast cancer mutation data base. http://research.nhgri.nih.gov/bic/?CFID=313172&CFTOKEN=38988484.
[7] Peshkin BN, DeMarco TA, Brogan BM, Lerman C, Isaacs C. BRCA1/2 testing: complex themes in result interpretation. J Clin Oncol May 1, 2001;19(9):2555–65.
Current asthma prevalence percent by age. United States: National Health Interview Survey (NHIS) Data; 2010.
[48] Li JZ, Absher DM, Tang H, et al. Worldwide human relationships inferred from genome-wide patterns of variation. Science February 22, 2008;319(5866):1100–4.

[49] Raby BA, Silverman EK, Kwiatkowski DJ, Lange C, Lazarus R, Weiss ST. ADAM33 polymorphisms and phenotype associations in childhood asthma. J Allergy Clin Immunol June 2004;113(6):1071–8.

[50] Qu S, Sun D, Wang Y, Zhang C, Lv Y, Yao L. Association of ADAM33 polymorphisms with childhood asthma in a northern Chinese population. Exp Mol Pathol December 2011;91(3):775–9.

[51] Van Eerdewegh P, Little RD, Dupuis J, et al. Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness. Nature July 25, 2002;418(6896):426–30.