



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.

# Public Health, Communicable Diseases and Global Health

## OUTLINE

<b>Public Health</b>	<b>169</b>	<b>Global Health</b>	<b>194</b>
<i>Introduction</i>	169	<i>Non-Communicable Diseases</i>	194
<i>Preventive Medicine</i>	170	<i>Obesity</i>	195
<i>Population Screening</i>	170	<i>Nutrigenetics and Nutrigenomics</i>	197
<i>Changing Behavior</i>	174	<i>Bioeconomy</i>	199
<i>Workplace</i>	175	<b>References</b>	<b>199</b>
<b>Communicable Diseases</b>	<b>178</b>		
<i>Detection</i>	178		
<i>Pathogenesis</i>	185		
<i>Emerging and Re-Emerging Infections</i>	192		

## PUBLIC HEALTH

### Introduction

*Public health* is a community-based strategy to improve health and well being, and to prevent disease through research, policy, education and appropriate practice. It is very different to *personalized medicine* which focuses on the individual. Common to both is the potential for DNA-based information to enhance clinical care. Fundamental to the practice of public health is *epidemiology* – the study of the causes, distribution, control and prevention of diseases

within a population. Until fairly recently epidemiologists relied on traditional approaches and measures of population health. Now an additional dimension is available through molecular (DNA) testing.

The potential to study the interactions between genes and environments is a powerful instrument for those whose research and clinical focus is a population. Tools used by geneticists, such as genome wide association studies (GWAS) and DNA banks are now accepted as legitimate methodologies for research undertaken by public health professionals and epidemiologists. A flavor of what is possible in

public health genomics can be found in the Centers for Disease Control and Prevention website [1].

## Preventive Medicine

The concept of prevention is a gold standard in public health, moving the focus from treating an established disease to maintaining well being, and avoiding disease or delaying its onset. As well as research and education, prevention requires appropriate *interventions*. In the prevention of a disease such as cervical cancer, screening is a core activity, but there is another preventive focus which is directed at finding *risk factors*, for example elevated blood cholesterol. The *social determinants* of disease, particularly in developing countries, and the underprivileged, under-resourced or minority groups are also being emphasized and factored into prevention strategies.

There are multiple preventive approaches (Table 6.1), and prevention is not all-or-nothing as interventions are possible during different phases of pre-disease and disease development (Figure 6.1). This expanding view of prevention is considered by some to be a weakening of the concept, and even the terminology is confusing with *public health* and *population health* used by some to mean the same thing (the pragmatic approach adopted in *Molecular Medicine*) while others distinguish the two [2]. In this complex mix, the application of molecular (DNA) knowledge provides additional options for prevention strategies; from the earliest possible detection of disease development to novel therapies.

## Population Screening

There are different ways DNA genetic testing can be incorporated into population screening programs once established criteria balancing risk versus benefit are adequately addressed (Table 6.2). Screening programs developed through a public health perspective have as

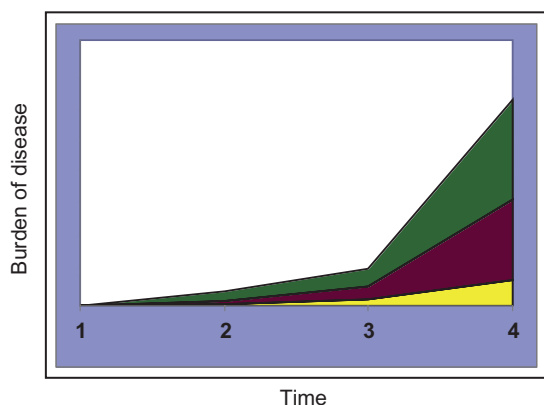
the primary focus the *health of the community*, while screening programs coming from a genetics viewpoint consider the *individual's rights* to be paramount. Hence, the philosophy behind the consent process can be different. This is illustrated by the newborn screening program.

From a public health perspective, screening newborns to prevent a serious disorder such as congenital hypothyroidism, with its associated severe intellectual impairment, can lead to important clinical outcomes. The screening test itself poses no risk to the newborn (compared to, for example, some vaccinations) and the benefits are significant. Therefore, what type of consent process is needed? The options vary from no consent (if newborn screening is mandated by law) to an opt-out consent process to a fully informed written consent (more

**TABLE 6.1** Preventive measures can be implemented at different stages of disease [2].

Prevention	Description	Examples
Primary	Promoting health prior to development of disease or injuries	Immunization, health promotion campaigns including anti-smoking and healthy diet choices.
Secondary	Detecting disease in its early (asymptomatic) stages	Screening, case finding, early detection.
Tertiary	Reversing, arresting or delaying progression of disease	Preventing complications of chronic diseases such as diabetes including rehabilitation.
Quaternary	Avoiding consequences related to overmedication, overdiagnosis or incidental findings, e.g. imaging.	Availability of medical based information from the Internet leading to an increase in the <i>worried well</i> . Direct-to-consumer DNA genetic testing (Chapter 5).

on consent in Chapter 10). Some argue that informed consent is essential because the parent(s) must be engaged to ensure they will provide regular supplementation for congenital hypothyroidism in an affected child. Others would say that the health implications of an affected child for parents and society make screening the highest priority. Less seems to be written about the child and who protects his or her right to a healthy life. Different arguments based on a public health versus a genetics approach are possible [4].



**FIGURE 6.1** Progression of disease with burden (Y axis) plotted against time (X axis). Three different disorders (green, red and yellow) develop over four time periods with gradual increase in burden of disease over this time. Ideally, detection of disease onset should occur at the earliest phase (1) rather than waiting until the burden is such that treatment becomes very difficult (phase 4). Some disorders lead to significant burden of disease (green) while others move along at a relatively lower level of activity (yellow). Preventive steps should be available for all four phases with the goal being to push back from 4 to 3 to 2 and ideally 1 to optimize treatment outcomes. DNA testing (screening) plays an important part in this push-back as it has the potential to detect genetic mutations that predispose to disease (germline DNA) or early signs of disease perhaps in somatic cell DNA in tumors (phases 2 or 3). While the focus is on DNA one should not forget that developments in other omics (particularly proteomics, metabolomics or epigenomics) can identify changes that assist earlier diagnosis.

The types of screening programs available through DNA genetic testing are listed in Table 6.3. The utility of DNA-based strategies, particularly the potential for PCR to test many samples quickly and cheaply, has meant that widespread screening of a population becomes a practical consideration. As more genes are sequenced, the number of mutations identifiable by PCR will increase. Omics-based technologies will continue to expand the options for DNA testing.

Two early examples of selective screening programs targeted to at-risk populations illustrate some advantages and disadvantages of this type of testing. Tay Sachs disease is a fatal neurodegenerative disorder of childhood. It is inherited as an autosomal recessive trait. Since the early 1970s, individuals at risk of having

**TABLE 6.2** Some criteria that a disease should meet prior to it being considered suitable for population screening [3].

Category	WHO principles of early disease detection
Condition	<ul style="list-style-type: none"> <li>• It should be an important health problem</li> <li>• There should be a recognizable latent or early symptomatic stage</li> <li>• The natural history should be adequately understood</li> </ul>
Test	<ul style="list-style-type: none"> <li>• There should be a suitable test (see [3] and Table 3.6 for criteria on DNA tests including sensitivity, specificity and so on).</li> <li>• The test should be acceptable to the population</li> </ul>
Treatment	<ul style="list-style-type: none"> <li>• There should be an accepted treatment for patients with the disease</li> </ul>
Screening program	<ul style="list-style-type: none"> <li>• There should be an agreed policy on whom to treat</li> <li>• Facilities for diagnosis and treatment should be available</li> <li>• The cost of case-finders should be economically balanced in relation to possible expenditure on medical care as a whole</li> <li>• Case-finders should be a continuing process</li> </ul>

**TABLE 6.3** Some examples of DNA-based screening strategies.

Screen	Explanation
Family screening	Family members at increased risk for a genetic disorder can be screened and this information used for early interventions or for decisions in reproduction. This approach requires the family-specific DNA mutation to be known. For example, cystic fibrosis screening in parents and siblings can be implemented if a newborn child is shown to be heterozygous for the p.Phe508del mutation. Family screening has also been called <i>cascade testing</i> because once it is established which side of the family has the particular risk allele, screening via family tracing can be undertaken.
Population screening	(1) Pre-disease DNA testing – a contemporary DNA population screening dilemma is hemochromatosis. (2) Carrier testing for single gene Mendelian disorders such as Tay Sachs disease in Ashkenazi Jewish populations.
Newborn screening	An accepted approach to test for reversible or treatable genetic or congenital disorders in newborns without any specific risk. Does this by analyzing a drop of blood collected by heel prick.
Workplace screening	The options available are discussed in the text.

children with Tay Sachs have had the opportunity for genetic screening and counseling. As a result, the incidence of Tay Sachs disease has been reduced without the societal problems that developed following the implementation of population screening for sickle cell disease.

Sickle cell disease is an autosomal recessive disorder with around 100 million carriers worldwide, and 2 million in the USA, most of whom are African Americans. There can be considerable morbidity and mortality associated with those who are homozygous affected, although the ultimate outcome is not entirely genetic in origin, as environmental factors are important

(more on this in Chapter 2). The US-based sickle cell screening program, which was also started in the early 1970s, targeted the at-risk African American population. The initial version of this program produced more harm than good. Results led to a lowering of self-esteem, over-protection by parents and discrimination. The discrimination came from employers, insurance companies, health insurers and potential spouses. Why did the two screening programs produce different outcomes?

One reason for successful Tay Sachs screening was the nature of the target group, which comprised individuals of Jewish origin who had better educational opportunities and social infrastructure. Another contrast between the two programs was the close community consultation undertaken prior to testing for Tay Sachs. Because of the problems associated with sickle cell screening, changes were made, including the removal of legal compulsion to be screened and improved counseling and education facilities. These enabled more successful testing to be pursued. Experiences with these programs illustrate the necessity for counseling and public education to explain the significance of mass screening results as key ethical considerations in design. Today, there are other population DNA screening dilemmas including: (1) Cystic fibrosis population screening, and (2) Sickle cell trait screening in sport.

### **Cystic Fibrosis Population Screening**

Over 1500 mutations in the *CFTR* gene produce cystic fibrosis, although p.Phe508del is the most common found in northern Europeans. Others are much less frequent. So how useful is a test that will not detect all those who are affected? For example, if only the p.Phe508del mutation is sought, false negative results in couples from a population with a frequency for this mutation of 70% will be  $0.51 (1 - (0.7 \times 0.7))$  – i.e. approximately half the couples will not be identified by this approach. Detection of the less

common mutations (some of which are only present at a 1–2% frequency in the population) would add to the workload, but would not substantially increase the information gained by the screening program. Additional problems that would need to be resolved before embarking on widespread cystic fibrosis screening include:

1. Uncertainty about disease severity for some mutations. Thus, counseling in a number of instances will be difficult and incomplete, and
2. Potential for racial profiling as cystic fibrosis is rarely found in some populations, for example Asian, and so detection rates in these will be minimal. Some would argue that one should not exclude or include particular ethnic groups in screening programs because this places undue emphasis on ethnicity and predisposition to genetic diseases. Others would say that disease and ethnic predisposition is a reality and, in the context of personalized healthcare, needs to be considered.

Debate continues about the value of cystic fibrosis mass population screening in contrast to testing individuals or at-risk families (selective screening). Even if laboratory facilities were available, major genetic counseling and public education efforts would be required to ensure that those tested fully understood the implications of the results. The financial resources needed to carry out a mass screening program would be enormous. In view of this, and the inability to detect all mutations with present technology, recommendations vary. In the USA the recommendation is for limited screening – perhaps of pregnant women, or selective screening of groups or families who are at higher risk than the general population [5]. Other countries do not recommend screening of pregnant women. A 2010 European consensus statement on carrier screening provides an in depth overview as well as a framework for what might be possible in member states [6].

### ***Sickle Cell Screening in Sport***

Population screening for sickle cell disease is in place in some newborn screening programs, particularly those involving at-risk populations. Sickle cell disease has potentially fatal consequences, but its effects can be ameliorated or avoided by early medical intervention including the use of antibiotics. DNA testing for the sickle cell *trait* is used in at-risk couples or populations, since offspring of an at-risk couple have a 1 in 4 chance of inheriting sickle cell disease.

As will be discussed below under Workplace, DNA testing can be used to screen selected populations to detect individuals who are at risk of a work-related illness. In this context, work can include sport. Since hypoxia is one precipitant for an acute attack in sickle cell disease, one might see justification in screening players involved in a sport likely to lead to hypoxia. What to do with this information could be problematic, but the issue is already facing some sporting bodies, as exemplified by the case of a 19 year old university student who died as a result of a rare complication of sickle cell *trait*, and the subsequent court action. In this case, the organization responsible for student sports at this level determined that sickle cell trait screening would become mandatory despite the trait, in contrast to the disease, rarely leading to serious medical complications [7].

Screening for a *trait* is another example of the public health versus the genetic approach, with the latter considering sickle cell trait to be a *good* trait since it has evolved with time to protect against malaria. Therefore, care is taken to avoid discrimination against or stigmatization of carriers. In contrast, the public health (or more likely in this case the medico-legal) perspective views the trait as a risk factor that needs to be screened for, to identify those who might need appropriate interventions or, more problematic, exclusion from a sport. It will be interesting to see how this controversial screening program for an autosomal recessive *trait* unfolds.

### **Newborn Screening**

Taking blood from the newborn's heel to test for *treatable* and/or *preventable* medical disorders has been in place since the early 1960s. Initially this was undertaken with biochemical testing and then DNA analysis was added. Next, tandem mass spectrometry (Chapter 4) became possible, allowing metabolomic-type approaches to screening for amino acids, organic acids and fatty acid metabolism to be included [8].

Today, there is little dispute that screening newborns for treatable disorders such as phenylketonuria and congenital hypothyroidism are important public health initiatives. Less clear is the value of newborn screening for a variety of other conditions, including the hemoglobinopathies, galactosemia, maple syrup urine disease, homocystinuria, biotinidase deficiency, congenital adrenal hyperplasia and cystic fibrosis [5]. The options for screening have been further expanded by tandem mass spectrometry, with its potential to detect many metabolites both normal and abnormal [8]. The former is an important consideration, since false positive results from screening will place additional pressure on the health system as well as increasing the *worried well* (Table 6.1). The debate about informed consent, presumed consent or even legal compulsion in public health measures such as newborn screening will continue for some time.

### **Changing Behavior**

The applications of molecular medicine in public health practice have introduced new options for preventive programs and interventions. However, changes will only occur if health professionals (starting with medical students) understand the implications and basis for molecular medicine and incorporate this knowledge into their work.

Will DNA based knowledge lead to better health choices by members of the community? Data on this are only now starting to be

gathered. One review found evidence that DNA genetic testing for rare genetic variants such as the *BRCA1* and *BRCA2* genes in breast cancer does lead to changes, such as follow-up mammograms [9]. Less clear was whether this knowledge influenced the behavior of other at-risk family members. The health literacy of the population remains a critical factor in whether behaviors change. If so, statistics emerging from the same review are worrying; more than a third of US adults have limited health literacy and only about 12% have sufficient health literacy skills to understand this type of information [9] (see Chapter 10 for more discussion of education).

### **Familial Hypercholesterolemia**

It is worthwhile concluding this section with a scenario discussing *from the laboratory to the bedside*, although in today's philosophy of avoiding hospitalization and expensive medical interventions we should be saying *from the laboratory to the community*. The example is familial hypercholesterolemia (FH), an autosomal dominant Mendelian disorder which is reasonably common in many populations, affecting about 1 in 500 people in a country like the UK. Familial hypercholesterolemia is clinically important, as >50% of affected men will develop coronary artery disease by the age of 50, and >30% of women will do so by the age of 60 [10]. Heart UK also estimates that of the 120 000 predicted to be affected in the UK, only 15 000 have been identified [11]. Can public health measures utilizing DNA testing help to bridge this gap? Presently the standard criteria of family history, clinical examination and serum cholesterol measurement are insufficient, particularly if familial hypercholesterolemia needs to be detected earlier to optimize the effect of anti-cholesterol drug therapy.

Our molecular understanding of familial hypercholesterolemia started in 1972, when M. Brown and J. Goldstein used biochemical and cell culture approaches to study this disorder. Subsequently they showed that cholesterol

metabolism was controlled by a receptor called LDL (low density lipoprotein) and abnormalities in it would lead to familial hypercholesterolemia. They were awarded the Nobel Prize in Physiology or Medicine in 1985 for their work. Once the *LDLR* gene for this disorder was isolated, DNA tests for a variety of purposes (diagnosis, prediction and screening) could be developed.

The addition of DNA testing in the management of familial hypercholesterolemia now improves the diagnostic accuracy, and the same test can be used to identify at-risk family members. However, this comes at a cost. DNA testing is not simple, as the *LDLR* gene is large and mutations are often family-specific. Therefore, DNA sequencing is needed and any changes found are not necessarily pathogenic in nature, but can be variants of unknown significance (Chapter 3). Mutations in other genes can also produce a similar clinical picture (phenotype). These include *APOB*, *ARH* and *PCSK9* which interfere with the cholesterol pathway. Finally, environmental factors such as diet, smoking and hormones also impact on the cholesterol level. Thus, the costs and considerable work involved would need to be balanced against the clinical benefits of earlier diagnosis for individuals, families and the broader community.

Failure to make a diagnosis of familial hypercholesterolemia might have been less of an issue before cholesterol-lowering drugs such as the statins became available. Today, treating an individual with elevated cholesterol is very effective, and it is generally believed that intervening early avoids cardiovascular and related complications of familial hypercholesterolemia. In 2008 NICE (the UK's National Institute for Health and Clinical Excellence) published guidelines for a new approach to the treatment and diagnosis of familial hypercholesterolemia, which included personalized medicine through DNA testing of individuals and at-risk family members detected by cascade testing. In the Netherlands there is an

ongoing, community-based, familial hypercholesterolemia screening service run by specialized nurses. It has produced some impressive detection rates which are expected to reduce morbidity and mortality in the longer term. The NICE guidelines allow a similar approach in other countries. It will be important to evaluate the clinical effectiveness of this preventive measure utilizing DNA testing.

## Workplace

DNA testing in the workplace could be undertaken for:

1. Detecting predisposition to disease or injury because of genetic susceptibility;
2. Detecting exposure to toxins;
3. Litigation, and
4. Identity checks [12].

### ***Detecting Predisposition to Disease or Injury***

This is the most contentious of the four applications, since it implies that DNA genetic testing can predict who will develop an illness or an injury in a particular work environment. One example of the approach is beryllium exposure, which occurs in industries such as defense, aerospace, nuclear power, electronics and dental prostheses. Even if a worker is not directly dealing with beryllium, secondary exposure can occur via airborne particles. Family members exposed to dust carried on clothing or footwear may also be at risk. Individuals sensitized to beryllium are at risk of developing acute or chronic disorders of the skin and lung, with the most serious consequences being carcinoma of the lung or chronic granulomatous lung disease (chronic beryllium disease).

Research has shown that genetic variants of the *HLA-DPB1* gene, particularly *HLA-DPB1<sup>E69</sup>* are found more often in exposed workers who go on to develop a cell-mediated, type IV,



delayed hypersensitivity reaction, leading to chronic beryllium lung disease. Mortality associated with this complication is around 36–62% [13]. However, it is important to note that the HLA genotype *per se* is insufficient to lead to disease and within the environment there are modifying factors such as the type of job; e.g. machining is more risky.

Will testing for *HLA-DPB1* variants predict which workers are likely to develop beryllium related disease? Despite the odds of lung disease associated with the glutamic acid 69 variant being high (84% of workers with chronic beryllium disease versus 36% in exposed workers without this disorder), the DNA test would not be particularly helpful, because the prevalence of *HLA-DPB1<sup>Glu69</sup>* in the normal population is high (40%) while the prevalence of disease among beryllium workers is relatively low (5%)

so the positive predictive value of 11.7% is not high enough to make DNA testing a worthwhile screen [13].

Other examples highlighting ethical and legal dilemmas include the *APOE4* DNA marker and predisposition to dementia following head injury in boxing (Box 6.1). Another genetic link between sport and illness is autosomal dominant familial hypertrophic cardiomyopathy, which is caused by mutations in muscle sarcomere genes. This disorder may initially present as sudden cardiac death following strenuous physical activity. Although the molecular DNA defects underlying this disorder are known, their number and complexity make it impractical to screen professional sportsmen and women, unless there are reasons such as a family history, unexplained syncopal attacks, or cardiac findings during clinical examination. Generally, an individual with

### BOX 6.1

#### GENES AND SPORT.

The *APOE4* gene variant described earlier (Chapter 2) is associated with a greater risk of developing Alzheimer disease, and the risk appears to be further increased in boxers – presumably as a consequence of chronic brain trauma. In a recently reported study, 50% of individuals with chronic traumatic encephalopathy were shown to carry at least one *APOE4* allele (one was homozygous for this marker) compared to the general population carrier rate of 15% [14]. Although considering only a small sample size, a 2006 report suggested that the *APOE4* variant was also associated with poorer cognitive and behavioral outcomes following moderate and severe traumatic brain injury [14]. Should an individual who has the *APOE4* marker (particularly someone who is homozygous for this marker) avoid boxing? Would an employer or

trainer be at risk of litigation for not advising a boxer to have their *APOE4* status determined? Should someone with this genetic marker be excluded from boxing? Hypothetical questions such as these continue to be asked, but there are no clear answers. If genetic testing is used for screening for susceptibility to work related conditions it should show:

1. Strong evidence for linking the working environment and the disorder;
2. The disorder has serious implications for the health or safety of employees;
3. The test has the appropriate sensitivity, specificity and other parameters, and
4. Privacy and the potential for inappropriate discrimination are addressed.

this type of inherited cardiomyopathy is warned against playing competitive sports as strenuous activity is associated with sudden cardiac death. Those with the disorder can have their heart rates monitored electronically, or have defibrillators implanted to instantly revert ventricular arrhythmias that arise.

### **Detecting Exposure to Toxins**

There are many potential toxins in the workplace. Genetic monitoring has been used in circumstances involving radiation and genotoxic chemical exposures. Detecting damage to DNA is important but difficult, especially at low exposures where health effects may not become apparent until well into the future. As was shown after the Chernobyl nuclear power reactor accident in 1986, chromosomal damage in workers exposed to significant  $\gamma$  radiation in the clean-up operation was an important indicator of damage. However, age and smoking habits were confounding factors for genetic damage, and the costs of FISH assays for detecting chromosomal abnormalities were too high for large scale population studies [12].

A new approach to detecting DNA damage might be possible with Next Generation (NG) DNA sequencing, which is interesting since detecting radiation-induced DNA damage was one of the early reasons for initiating the Human Genome Project (Box 1.2). The potential for quantitating cellular and tissue damage is illustrated by the use of this to study genomes of patients with lung cancer caused by cigarette smoking. Tobacco smoke contains more than 60 carcinogens, and damage results from chemical modification of purines by mutagens, inability of the DNA repair mechanisms to correct this damage and incorrect nucleotide incorporation opposite the distorted base during DNA replication [15]. NG DNA sequencing allows the DNA signatures of tissue damage and DNA repair to be cataloged. It may show sufficient specificity to permit monitoring of the environment (by screening workers) or detect when damage

has been caused and by what particular toxin (screening workers with illness).

### **Litigation**

Quantifying the evidence of exposure is a significant hurdle in a tort action (called toxic tort if the wrongful act involves exposure to a toxic substance). It is not easy for a plaintiff to prove that exposure to a toxic substance has occurred *and* that the toxic substance was the cause of illness or injury. Conversely, a defendant in a toxic tort may have difficulty disproving a claim because of doubtful or minimal evidence. However, exposure to *xenobiotics* (compounds that are foreign to the body) will provoke changes in gene expression in any biological system. This is the rationale behind the use of transcriptomics to identify or characterize changes that result from exposure to toxins. There is potential for *toxicogenomics* to provide a new and more definitive evidence of exposure to a toxic substance by looking for particular cellular responses before and after exposure to it.

### **Identity**

Workplace DNA testing to establish identity is used in the military and the police. The purpose is to have on record a reference DNA profile for identifying, if necessary, body parts (war, fighting or terrorism) or to assess crime scene contamination. These aims are not controversial but concerns include:

1. Security of the DNA sample, and who has access to it;
2. Will the DNA sample, or more likely the DNA profile, be included in the databases which are used to search for criminal activities?
3. How long are the DNA samples/profiles kept – i.e. are they destroyed once the individual is no longer in the military/police, and
4. Is this a voluntary or compulsory part of the employment agreement?

## COMMUNICABLE DISEASES

### Detection

There are many applications of molecular medicine in the communicable diseases caused by bacteria, viruses, fungi, parasites and in a rare example by an abnormal protein. As well as the known infectious agents, there are the newly emerging (or re-emerging) infections and an increasing number of immunocompromized patients exist. To this mix the development of therapy-resistant organisms and bioterror can be added. In such a changing environment, no single therapeutic or preventive approach will be sufficient. What is certain is the ongoing requirement for rapid and accurate detection of infectious agents, which is best undertaken by molecular-based diagnostics. In infectious diseases, these are usually known as NAT (nucleic

acid testing) because they involve both DNA and RNA.

Previous editions of *Molecular Medicine* gave an in depth overview of how knowledge of DNA could be used to improve the detection of infectious agents for patient care. This detail is no longer necessary because DNA testing is now used routinely in clinical management and public health strategies. The various diagnostic tests derived from the traditional phenotypic tests to DNA-based genotypic tests are summarized in [Table 6.4](#).

As already noted, the utility of DNA sequencing, particularly for viral infections (because their genomes are relatively small), has expanded rapidly and now contributes key data for investigating new outbreaks. Just as occurred in genetics, an omics approach will become increasingly preferred – already the concept of *infectomics* is being touted. More

**TABLE 6.4** Two approaches to laboratory testing in microbiology are the traditional phenotypic tests or the new genotypic DNA or RNA tests (NATs).

Phenotype-based tests	Genotype-based tests
Traditional diagnostic approaches include:	Strategies for analyzing pathogen nucleic acid tests:
<ul style="list-style-type: none"> <li>• Microscopy – staining, appearance</li> <li>• Culture and growth characterization</li> <li>• Biochemical testing</li> <li>• Immunological profiling (antisera or antibodies).</li> </ul>	<ul style="list-style-type: none"> <li>• Nucleic acid hybridization</li> <li>• Plasmid identification</li> <li>• Chromosomal DNA banding patterns</li> <li>• PCR amplification techniques</li> <li>• Microarray based assays</li> <li>• DNA Sequencing.</li> </ul>
Can provide clues for identifying new pathogens. Tried and trusted approaches that are often relatively cheap and technically easier than genotypic methods. However, can be slow and so not always useful during epidemics, emergencies or new infections.	Like the trend in genetic disorders, DNA sequencing is assuming greater utility for detecting infectious agents [16]. Unlike genetic DNA testing, contamination is a major source of error because there is considerably less template DNA.
Phenotypic variation can occur during pathogens' life cycles making it difficult to interpret results at times.	Variation less of an issue but finding DNA or RNA does not necessarily confirm an organism to be pathogenic. For example, the detection of CMV DNA by PCR in a patient's serum could mean active disease or latent infection.
Host immune responses can be delayed or may remain persistent even after resolution of infection. Cross-reacting antibodies from natural infection or vaccination can produce false positive results.	Best for detecting difficult to culture organisms or there is a mix of pathogens. DNA testing has greater sensitivity and also allows virulence factors and drug resistance to be detected. Q-PCR helpful in monitoring treatment with viruses such as HCV and HIV.

sophisticated bioinformatics is being developed to deal with metagenomics (Chapter 4) and this will ensure that new software will allow the sequence information from complex mixes of organisms (even those in clinical specimens) to be analyzed and separated into distinct organisms.

Evaluating a NAT is based on traditional measures:

1. Sensitivity;
2. Specificity;
3. Positive predictive value (PPV), and
4. Negative predictive value (NPV) (Table 3.6).

Tests with high PPVs are needed for infections where a false negative will have significant clinical or psychological consequences, for example, tests for sexually transmitted infections. Tests with high NPVs are required when it is essential that positives are not missed, for example blood screening.

### **Taxonomy and Comparative Genomics**

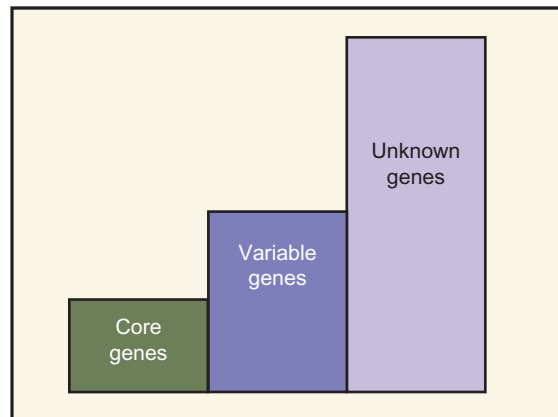
The first microorganism to be sequenced was *H. influenzae* in 1995. Since, there have been large numbers of microbial and viral sequences deposited in databases, including both pathogens and non-pathogens. Completed, whole-genome sequences exist for around 3000 bacteria, 41 eukaryotes (19 of these being fungi) and 2675 viruses. In addition, 40000 and 300000 partial sequences for influenza and HIV-1, respectively, have been completed [16]. The numbers of sequenced microorganisms will continue to grow exponentially and metagenomic approaches will allow the detection of many novel organisms (Chapter 4). The larger databases available for study will ensure sophisticated comparative genomics can be undertaken for research and clinical applications. DNA-based information is adding a new dimension to taxonomic classification, as described below for viruses.

As multiple *de novo* sequences of the same organism are obtained, it has become

apparent that there is a *pangenome*. This means that different strains of an organism have:

1. The same core genes;
2. A number of genes that are variable and used for adaptation to particular environments, and
3. A set of genes with no known function (Figure 6.2).

The pangenome varies between organisms, for example, all genes for *B. anthracis* appear to be present in only four species. In contrast, for *E. coli*, it is likely that the pangenome will require hundreds of these bacteria to be sequenced. Apart from providing further insights into the structure and function of organisms, knowledge of the pangenome is likely to be more informative than any individual genome when considering new virulent forms or the development of drug resistance. To study and understand the pangenome requires an omics approach. It is also apparent that while microbial genomes are small compared to eukaryotes (Table 1.7) they are relatively rich



**FIGURE 6.2** The pangenome comprises all genomes in a group of organisms [16,17]. The pangenome is divided into: (1) Core genes – essential for basic function; (2) Variable genes – these reflect the environment that the organism needs to deal with, and (3) Unknown genes – found on DNA sequencing but function is unknown. The relative sizes are not drawn to scale but are meant to show a smaller core, with large numbers of genes with unknown function.

in protein-coding genes (humans 1–2%, microbes 90%) [16,17].

Unlike all other cells that have DNA as their genetic material, viruses are considerably more diverse in what they use. This is reflected in a molecular classification that defines seven different viral classes on the basis of their genetic material and replication strategies. ds – double stranded; ss – single stranded; (+) – positive-sense or plus strand; (–) – negative-sense or complementary strand:

- dsDNA – example is adenovirus
- ssDNA – adeno-associated virus
- dsRNA – rotavirus
- ss(+) RNA – poliovirus
- ss(–) RNA – rabies virus
- ss(+) RNA plus reverse transcriptase – retrovirus
- DNA plus reverse transcriptase – HBV.

Viruses are the smallest organisms, and have genome sizes measured in kilobases. The International Committee on Taxonomy of Viruses (ICTV) develops an agreed taxonomy and nomenclature. It maintains an official index and publishes this information. In its 2009 release, the ICTV recognized six orders of viruses with another group yet to be placed into an order. There were 87 families, 19 subfamilies, 348 genera and 2285 species confirming further the heterogeneity found in viruses. The building of an accurate taxonomic classification has many advantages, including new insights into the biology of the viruses and their evolutionary relatedness which provide important clues when dealing with new infections [18] (Box 6.2).

Applications of DNA sequencing in virology include:

1. Identifying the function of viral proteins to allow a better understanding of how viruses evade host immune responses or promote their own migration and spread;
2. Defining regulatory controls or proteins that might become targets for new anti-virals;

3. Developing rapid diagnostics and detecting the identity of new viral outbreaks, and
4. Understanding evolution and hence relatedness for molecular epidemiologic strategies investigating outbreaks of old and new viruses, and monitoring drug resistance [18].

### ***Nosocomial Infections and Drug Resistance***

Nosocomial, or hospital acquired, infections are usually associated with medical devices such as catheters, or surgical procedures. Apart from wound and urinary tract infections they lead to life-threatening pneumonia and septicemia. Some statistics on these types of infections include:

1. They were the sixth leading cause of death in the USA in 2002 with approximately 99 000 deaths;
2. Estimated cost to the US Healthcare budget is over \$5 billion annually;
3. Approximately one third are preventable, and
4. Gram negative bacteria are involved in more than 30% of infections [20].

The convergence of gram negative bacteria that are increasingly antibiotic resistant and a reduction in drug development programs has produced a gloomy scenario for hospital acquired infections.

Causes for antibiotic resistance are many including:

1. Unnecessary or inappropriate use of antibiotics in humans;
2. Availability of antibiotics over the counter;
3. Use in the food industry including meat, agriculture, aquaculture;
4. Poor patient compliance in taking prescribed drugs;
5. Transmission by farm or pet animals treated with antibiotics, and
6. Inadequate infection control measures in hospital and clinical care (Table 6.5).

New drugs are not being developed as quickly as they are needed because of high production

**BOX 6.2****INTEGRATED DNA AND RNA IN THE GENOME.**

The genomes of vertebrates contain many copies of retroviral sequences acquired during evolution. These could function to protect the host from viral infection, and possibly as a source or natural reservoir for the virus to persist and transmit. However, it is now apparent that it is not only retroviruses that can integrate a copy of their RNA into the host's somatic and germline genome, which is the necessary first step before replication can occur. The genome of some bees has been shown to contain sequences from a positive (+) strand RNA *Dicistroviridae* that infects insects. These bees are resistant to infection by the virus. Following this observation, a comparative bioinformatics study of genomic sequences from 48 vertebrate species using sequence data from non-retroviruses containing single-stranded (ss) RNA genomes was undertaken. Surprisingly, it was shown that about half the vertebrates had integrated non-retrovirus sequences into their genomes. The next unexpected finding was that

these integrations came mostly from two groups of RNA viruses from the negative (-) strand RNA *Mononegavirales* order. These were either Ebola and Marburg viruses – *Filoviridae* family associated with lethal hemorrhagic fevers – or Bornavirus – *Bornaviridae* family are associated with neurological and psychiatric disorders which can be fatal. The vertebrates that had the integrations suggested these events had occurred over 40 million years ago. Therefore, the conservation of sequences coding for virus-like proteins is thought to have some selective advantage, possibly increasing the host's resistance to infection. Conversely, continued integration and persistence might provide viruses with a natural reservoir for future infections. An example would be bats, which are now thought to be natural reservoirs for the Ebola and Marburg viruses. Sequences from these viruses are detectable in some bats with some having open reading frames [19].

costs, the time required for clinical trials, regulatory demands and a concern that products will become obsolete once resistance develops. Apart from rapid diagnosis of the causative microorganism, improved detection of antibiotic resistance strains is also needed. These requirements can be met by a NAT approach, although this is only the first step of a more comprehensive internationally coordinated plan to address the issues of antibiotic resistance.

The urgency of this matter is well illustrated by tuberculosis (TB), where global control of this increasingly problematic public health challenge requires better and faster diagnosis of the primary infection as well as early detection of drug resistance. The traditional phenotypic

culture methods to diagnose TB are slow. Similarly, the first generation of molecular DNA diagnostic tests is complex, requiring sophisticated laboratory expertise and resources [26]. Multidrug resistant TB (defined as infections that are resistant to at least isoniazid and rifampicin) is emerging globally, particularly in India and China. Cases of extensively drug-resistant TB now exist, meaning that TB is also resistant to a number of the second line anti-TB drugs. Failure to detect resistant cases of TB is the rule rather than the exception, particularly where laboratory resources are limited. This means that new, DNA-based, detection kits, especially those that can be multiplexed and automated are eagerly awaited.

**TABLE 6.5** Some therapy resistant multidrug resistant organisms<sup>a</sup> [21–25].

Example	Type of resistance
Methicillin resistant <i>S. aureus</i> (MRSA)	Well known nosocomial infection and difficult to treat. Different MRSA strains are reported in relation to community-acquired infections usually in association with relatively minor skin or soft tissue infections. However, community-acquired MRSA can now lead to life threatening infections. Generally MRSA infections are spread through direct person-to-person contact hence many are preventable by hand washing. Both traditional culture and DNA testing can be used to detect MRSA. Which is preferred will depend on costs and laboratory staff skills.
Vancomycin resistant enterococci (VRE)	Unlike the higher profile infections caused by <i>S. aureus</i> , enterococci are less well known as nosocomial infections but can cause fatal diseases particularly if associated with vancomycin resistance (VRE). Individuals at risk are: (1) Hospitalized for a prolonged time; (2) Immunosuppressed, and (3) Post surgery or have devices such as urinary or intravenous catheters. Prior treatment with vancomycin and other antibiotics is an important predisposition. Resistance can occur intrinsically or acquired. Resistance genes include <i>vanA</i> , <i>vanB</i> , <i>vanD</i> , <i>vanE</i> and <i>vanG</i> with the first two commonly associated with VRE.
Gram negative bacilli	There is now antibiotic resistance emerging in the gram negative bacteria. Initially this appeared as plasmid encoded $\beta$ lactamases producing resistance to penicillin. It then expanded into ESBL (extended spectrum $\beta$ lactamases) producing resistance to penicillins, cephalosporins (1st to 3rd generations) and monobactams but not cephamycins or carbapenems. Today, there is added concern about the next trend involving NDM-1 (New Delhi metallo- $\beta$ -lactamase 1) because the carbapenem resistance gene ( <i>bla<sub>NDM-1</sub></i> ) has been detected by PCR. In the latter example, patients have acquired resistant <i>E. coli</i> or <i>K. pneumoniae</i> species in the Indian subcontinent and brought these back to the UK.
Tuberculosis (TB)	Multidrug resistance (MDR) TB involves the first line drugs particularly isoniazid and rifampicin (rifampin). Extensively drug resistant (XDR) TB is the next step with additional resistance including second line drugs.
Malaria	The single drug approach used initially to treat or prevent malaria has now given way to combination therapy including artemisinin (Box 4.6) as resistance emerges across the world. The molecular basis for drug resistance is complex involving many genes including pfCRT, pfMDR which is an ortholog of the P-glycoproteins found in mammals in association with multidrug resistance in cancer (Chapter 7), and mutations in the <i>DHFR</i> gene that produce resistance to antifolate drugs.
Influenza	Two mainstays for drug treatment during the influenza A (H1N1/09) virus pandemic that started in 2009 were the viral neuraminidase inhibitors oseltamivir (Tamiflu <sup>®</sup> ) and zanamivir (Relenza <sup>®</sup> ). A single amino acid mutation in H1N1 soon appeared (His274Tyr) and this produced a virus that was resistant to oseltamivir although the US Centers for Disease Control and Prevention indicated that almost all viruses remained susceptible. Examples of resistance to zanamivir were not reported.

<sup>a</sup>Generally these organisms are called MDROs (multiple drug resistant organisms) although some like MRSA or VRE are specifically referring to one type of antibiotic/drug resistance.

A fully automated NAT method to detect both TB and rifampicin resistance was reported in 2010. This uses uncultured sputum and can be completed in less than 2 hours with impressive sensitivities and specificities even in patients with TB and HIV, where smear-negative disease is more common. Since it is fully automated, it does not require sophisticated hands-on expertise. Although the NAT

only detects resistance to rifampicin it shows the way ahead, particularly if omics-based diagnostics including microarrays are developed [26].

### **Public Health Testing – Blood Transfusion Services**

Viruses such as HIV, HBV and HCV assume added notoriety when they are implicated in transfusion-derived infections involving blood

and plasma-derived products. Previously, blood transfusion services based their donor and blood screening programs on detecting antibodies or antigens in the donor or blood supply. However, this has proved to be inadequate, and an important addition to the screening protocols is the use of PCR to identify viral DNA or RNA. The advantages of a NAT include higher sensitivity and greater reliability during the window period – which is the time between a blood donor becoming infectious and donor screening tests becoming positive, i.e. seroconversion has occurred. The use of NAT, better serologic-based assays and more effective regulatory controls have made contemporary blood products considerably safer. Ultimately, transfusion services must balance safety against access to blood and its products. What is screened for will depend on the types of infections found within a geographic region as well as affordability of the screening tests.

NAT-based assays for screening blood donations can be used to screen pools of donations, for example, 16–24 donations simultaneously or individual ones. The former is more rapid and cheaper, but rare instances of HIV, HBV or HCV can be missed. The testing of individual donations is the method of choice but until recently was too expensive. Today, as new analytic platforms allow rapid and automated multiplexing NATs to be used, the screening of individual donations becomes more cost effective.

Blood transfusion services test blood and donors for a range of infectious agents depending on national requirements. The WHO recommends mandatory screening for HIV-1, HIV-2, HBV, HCV and syphilis, while the requirement for HTLV-I, HTLV-II (HTLV – human lymphotropic virus) and malaria are decided on a regional basis [27]. Other infectious agents that can be screened for include West Nile virus, dengue and emerging infections. Screening can also be undertaken in selected cases, for example, CMV free-blood for immunosuppressed

patients, the fetus or the neonate. The risk of prion diseases is considered below.

Ease of international travel means a potential donor could become infected elsewhere. This contingency is covered by donor questionnaires that allow self-exclusion (particularly for infections that are not routinely sought). For example, to prevent transmission of prion diseases through blood, some transfusion services have excluded donors who have lived in the UK over certain time periods (see below). Other reasons for deferral include fever with headaches the week before donation (a risk of West Nile and other viruses) or travel to certain regions (a risk of malaria).

### ***Awaiting Better Diagnostics – Prion Diseases***

A rare but important form of communicable dementia is found in the prion diseases (also called transmissible spongiform encephalopathy, or TSE). These diseases affect humans and a number of animals used for meat including cattle, deer, sheep and goats. The term *prion* comes from *protein* and *infectious* and was coined by S. Prusiner who was awarded the Nobel Prize in Physiology or Medicine in 1997 for his work on prions. The important components of prion disease include the *PRNP* gene and its cellular product PrP<sup>c</sup> (*prion protein cellular*) which can become the *infectious* protein product PrP<sup>Sc</sup> (*prion protein scrapie*). The normal PrP<sup>c</sup> is a cell surface glycoprotein found in a wide range of animals, having a function that as yet remains unknown. PrP<sup>c</sup> needs to change its conformation to its isoform PrP<sup>Sc</sup> to be infectious. No nucleic acid is involved in this process, highlighting the novel way in which prion disease arises and is propagated [28]. The disease leads to widespread neurodegeneration with cognitive and motor impairment. It is fatal and there is no treatment (Box 6.3). Work continues to develop an early diagnostic marker for this disease. This is a priority for screening blood and its products.

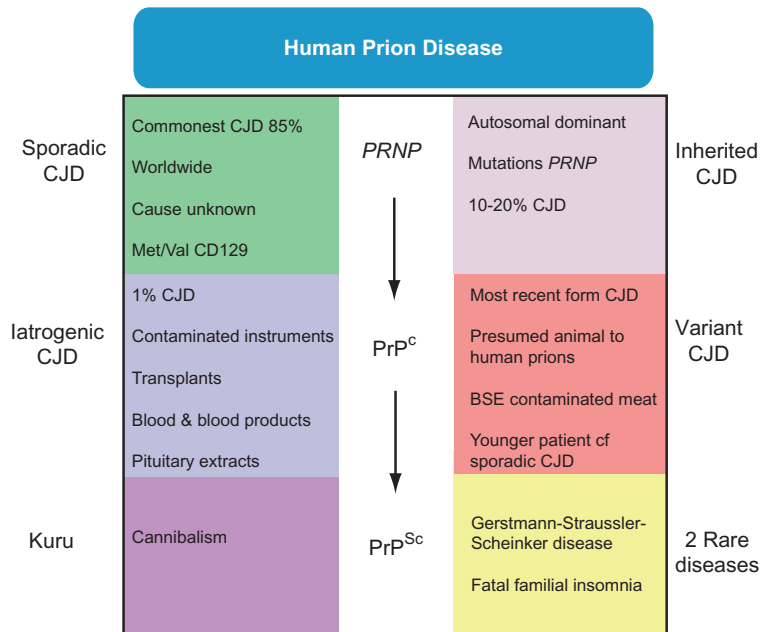


## BOX 6.3

## PRION DISEASE.

Prion disease may be sporadic, inherited, iatrogenic or transmissible from animal to human via infected meat and now human to human via blood products. The dementia that results includes sporadic, iatrogenic, inherited and variant Creutzfeldt-Jakob disease (CJD) in humans, bovine spongiform encephalopathy or BSE in cattle (related to the 1986 epidemic of *mad cow disease* in the United Kingdom), and scrapie in sheep and goats. In 1996, the emergence of variant CJD (vCJD) in humans is thought to have arisen from transmission across the species of the BSE agent. vCJD is characterized by

an early age of onset (Figure 6.3). Mutations in the *PRNP* gene account for the inherited forms of CJD. However, in the vast majority of sporadic cases, there are no detectable DNA mutations, and the change from PrP<sup>c</sup> to the abnormal PrP<sup>Sc</sup> is thought to occur because of somatic mutations or other, as yet unknown genetic or environmental factors. Risk factors for developing vCJD include young age, residence in the United Kingdom especially between 1985 and 1990, and intriguingly, homozygosity for a codon 129 polymorphism in the *PRNP* gene. At this position there is either a methionine or a



**FIGURE 6.3 Human prion diseases [28].** There are different types of Creutzfeldt-Jakob diseases (CJD), and two other prion-related diseases. The most recently described is vCJD which is thought to have occurred as a result of direct animal to human spread through contaminated beef products. Now there is evidence for human to human spread via transplants or blood products.

**BOX 6.3** (*cont'd*)

valine. In normal individuals, the combinations of methionine/methionine, methionine/valine and valine/valine are present. However, in patients with vCJD, the homozygous methionine is always found, suggesting that this may lead to genetic predisposition. Most patients developing iatrogenic CJD after receiving pituitary extracts for growth hormone are also homozygous for methionine. If this is correct, some have hypothesized that a second wave of vCJD will occur in the future involving those who are methionine/valine heterozygotes or homozygotes for the

valine allele because a longer incubation period is needed to develop prion disease without the additional genetic risk factor. Other less well characterized polymorphisms in this gene have been detected and may represent additional genetic modifiers [28]. Prion disease remains a challenge for the future, particularly to explain how the infectious forms occur without any apparent conventional infectious agents being involved. Better diagnostics and some form of therapy are needed for this rare but fatal infection.

## Pathogenesis

The pathogenesis of many infections has been determined from studies utilizing light/electron microscopy, cell culture or immunoassays. To these can now be added nucleic acid (DNA, RNA) based methodologies. Advantages provided by nucleic acid techniques include the ability to detect latent (non-replicating) viruses, and to localize their genomes to nuclear or cytoplasmic regions within cells. Tissue integrity remains preserved during *in situ* nucleic acid hybridization and so histological evaluation can also be undertaken. NAT can be manipulated to enable a broad spectrum of serotypes to be detectable. This is particularly valuable in emerging infections where the underlying serotypes are unknown. Today, a very powerful application of NAT is the ability to sequence whole genomes, and so identify a pathogen or what it is likely to be. From its genomic sequence it becomes possible to:

1. Predict its role in disease pathogenesis;
2. Find regions in the genome suitable for rapid diagnostics via NAT, and

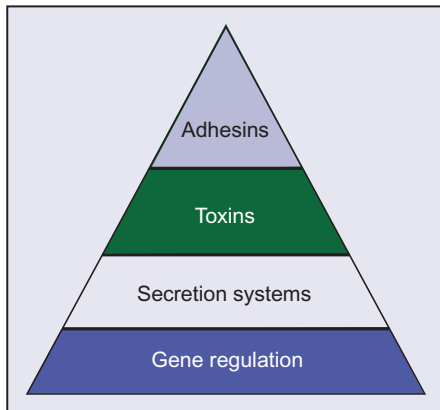
3. Consider how new treatment options including vaccines can be developed.

## Virulence Factors

Microorganisms have developed a range of virulence factors to allow them to invade a host (Figure 6.4). The best known are toxins, which are broadly divided into:

1. *Exotoxins* – usually proteins secreted by both gram positive and gram negative bacteria. They can be deadly, for example, tetanus exotoxin and diphtheria exotoxin, and
2. *Endotoxins* – usually heat stable lipopolysaccharides found in the gram negative bacterial cell wall.

Nevertheless, killing the host is not beneficial to the invading organism and in some circumstances it is essential that the host does not die. This is exemplified by *H. pylori*, which has sophisticated virulence factors including VacA and CagA allowing it to invade and cause damage to the host. However, the same organism has also evolved to ensure its continued survival by modulating its cell killing capacity because



**FIGURE 6.4 Virulence mechanisms used by bacteria.** Four mechanisms can be used by bacteria to invade a host. Which predominates will vary for each microorganism. (1) Adhesins allow bacteria to attach to host cells. This is the first step in the infective process. Some bacteria have appendages such as pili and flagella to facilitate attachment; (2) Many toxins are produced and have been well characterized both biochemically and molecularly; (3) Bacteria ultimately need to secrete their products into the host cell through specific secretory systems. A number have been described and are needle-like to allow the passage of toxins from the bacteria into the host, and (4) Implied in the concept of a pangenome is a complex bacterial genome to orchestrate the various changes needed to infect a host and produce the appropriate effects. The regulatory environment for this will need some common pathways and specific ones when comparing different bacterial species.

the CagA protein while cytotoxic *per se* counters some of the effects of the VacA toxin [29].

Toxins have many different actions, and using broad spectrum antimicrobials to inactivate them might not always suffice (Table 6.6). Nevertheless, the potential for this approach to treating or preventing infection is illustrated by *B. anthracis* – a bacteria causing anthrax. It achieved added notoriety because of an attempt at bioterror using postal letters in 2001 (Box 9.5). The attenuated anthrax bacteria (Pasteur strain) used for immunization lacks its toxin confirming the latter's importance in disease causation. Animal studies also suggest that antibodies that inhibit the anthrax toxin from binding to host receptors might provide

protection, at least in emergencies [30]. A better understanding of how toxins work and function as targets for new drugs is coming from molecular studies.

The traditional targets for conventional antimicrobials (usually antibiotics) include components of the bacteria that are essential for survival, such as the cell wall, the cell cycle, DNA replication and protein synthesis. This approach kills (bacteriocidal) or inhibits growth (bacteriostatic) of most bacteria, but invariably allows some residual subpopulations with natural immunity to be positively selected for, and hence the development of antimicrobial resistance will follow. Therefore, focus has now shifted to developing the next generation of antimicrobials, which target virulence factors. This would overcome the pathogenicity of the organisms without necessarily killing them and so avoids setting up an environment for resistance strains to emerge [30].

### Host Resistance

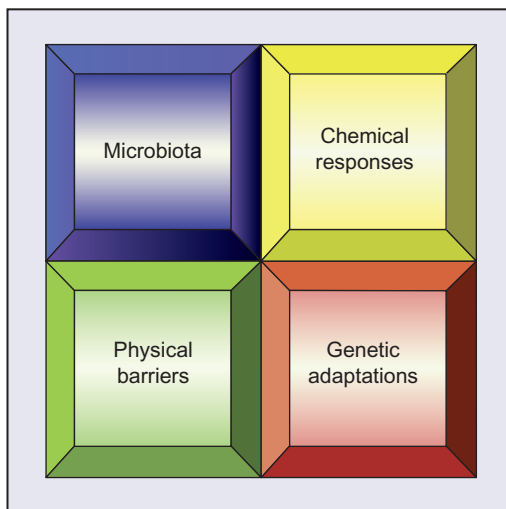
Microorganisms have developed sophisticated ways in which to invade a host, but hosts have evolved many protective mechanisms (Figure 6.5). The host's response in terms of genetic modifications is particularly relevant to molecular medicine. In humans, evidence for a genetic component influencing the outcome of an infectious disease comes from the following observations: (1) Not all exposed to HIV-1 get infected, and those who do progress to AIDS show different responses, and (2) Some ethnic groups are more resistant or susceptible to infections, e.g. resistance to malaria in some Black Africans.

**HIV-AIDS:** The main HIV co-receptor involved in the infection process is CCR5. Naturally occurring mutations in this receptor – such as a 32 base deletion present in up to 20% of European populations (about 1–2% are homozygous) – allow these individuals to be highly resistant (homozygotes) or partially resistant (heterozygotes) to HIV-1 infection

**TABLE 6.6** Some bacterial toxins in the gastrointestinal tract [30–32].

Bacterium	Toxins and their effects
<i>Clostridium botulinum</i>	Associated with foodborne illness. <sup>a</sup> Produces seven antigenically distinct neurotoxins that are important to detect. Conventional diagnostic assays are used although they are slow and difficult. A number of NAT have been developed and are being evaluated.
<i>Clostridium perfringens</i>	Associated with foodborne illness. Is a ubiquitous organism in nature. Produces two $\beta$ toxins detectable by traditional assays or PCR NATs.
<i>Escherichia coli</i>	The enterohemorrhagic <i>E. coli</i> (EHEC) remain an important cause of foodborne illness with one serotype 0157:H7 and other EHECs serious public health problems. Shiga 1 and Shiga 2 are the two main toxins and are so named because of similarities with the <i>Shigella dysenteriae</i> toxin. Molecularly the Shiga toxins have two subunits: A (active unit) and B (receptor binding unit). This toxin structure is similar to what is seen with the anthrax toxin although the latter has three subunits (1 for binding, 1 called the lethal factor and 1 called the edema factor). Rapid and sensitive methods to detect EHEC and its toxins for clinical purposes including source and spread are possible with NAT. In mid 2011, an outbreak of EHEC in Europe caused deaths and involved serotype 0104:H4. Its source was shown to be infected sprouted seeds. Using NG-DNA sequencing platforms, the genome for this pathogen was completed within a week. It showed the <i>E. coli</i> to be a hybrid strain and identified a number of antibiotic resistance genes. These findings might explain the pathogen's virulence and could also be used to design rapid NAT diagnostics.
<i>Vibrio cholerae</i>	There are 10 pathogenic vibrio bacteria associated with foodborne illness (particularly seafood) with cholera being the best known. PCR NATs have proven valuable in detecting the underlying vibrio as well as relevant toxins.
<i>Clostridium difficile</i>	A major cause of diarrhea in hospital patients and those in long term care. Serious infection is worsened by prior use of antibiotics that change the normal microbiota and allow proliferation of toxin producing <i>C. difficile</i> . A hypervirulent strain of this organism is spreading and is defined by NAT PCR as ribotype 027 which is thought to have risen by mutations in the toxin regulator gene leading to overproduction of toxins A and B. Its spread may be underestimated because NAT typing is not used in all countries.
<i>Listeria monocytogenes</i>	An important pathogen in the food industry with major outbreaks already reported in several countries. Virulence genes are located within a 9Kb cluster and are involved in ensuring cell to cell spread. They include a hemolysin gene ( <i>hlyA</i> ) with its product LLO essential for pathogenicity and three other genes. Detection methods include conventional agar plating but NATs provide greater flexibility particularly if large numbers of food products need to be screened.
<i>Helicobacter pylori</i>	Spiral organisms causing gastroduodenal disease including gastritis, peptic ulcer, gastric cancer and lymphoma. The importance of showing a link between these diseases and <i>H. pylori</i> was recognized with the award of a Nobel Prize for Physiology or Medicine to B Marshall and R Warren in 2005. Spreads from person to person and can produce a chronic life long infection unless treated. Non invasive but leads to chronic inflammation with cancer as possible sequelae. The two toxins are CagA and VacA and there are two types of <i>H. pylori</i> – 1 and 2. Each has the <i>vacA</i> gene but only type 1 has the <i>cagA</i> gene and so is the more pathogenic. Even though type 2 has the <i>vacA</i> gene it does not seem to express. There are many approaches to diagnosis including distinguishing types 1 and 2. NAT methods work well with gastric biopsies.

<sup>a</sup>Foodborne illness remains an important public health issue with major health and economic consequences. The US Centers for Disease Control and Prevention (CDC) estimates each year 1 in 6 Americans (48 million people) get a foodborne illness and around 3000 die.



**FIGURE 6.5 Host mechanisms to protect against invasion by microorganisms.** Various protective mechanisms allow the host to escape or modulate invasion by a microorganism. (1) Microbiota in the host (*microbiota* – normal microbial flora; *metagenome* (Chapter 4) – the genetic (DNA/RNA) material isolated from an uncultured microbial environment); (2) Physical barriers such as skin or mucosa, pH, temperature and secretions; (3) Chemical barriers particularly the immune response, and (4) Genetic adaptations which evolve over a long period of time but provide an effective mechanism to protect against certain pathogens.

and disease progression [33]. Studies are now underway with anti-HIV drugs that target the CCR5 receptor and a bone marrow transplant approach is described in Chapter 8.

**Malaria:** The two most common forms of malaria (*P. falciparum* and *P. vivax*) produce severe anemia. *P. falciparum* is also associated with cerebral malaria, respiratory and metabolic complications. This spectrum is partly explained by *P. falciparum* being able to invade a large proportion of red blood cells, whereas *P. vivax* can only invade the reticulocytes. Another explanation is the mode of entry of these parasites into red blood cells; *P. falciparum* has a number of routes of invasion, whereas *P. vivax* can only enter red blood cells that carry the Duffy blood group. This parasite is not seen in West Africa because the populations there are Duffy negative.

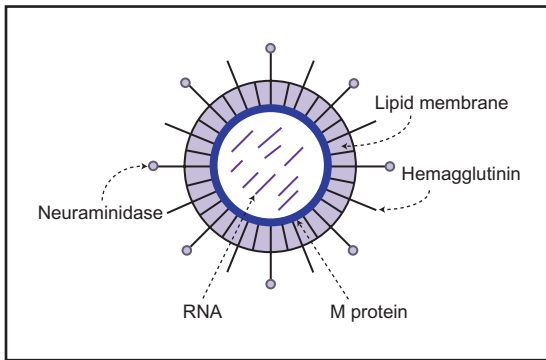
Host genetic factors that provide some protection from malaria have been identified. These include single gene effects seen in the hemoglobinopathies such as sickle mutation (HbS), HbE,  $\alpha$  thalassemias and  $\beta$  thalassemias. The hemoglobinopathy protective effect results from abnormal red blood cells that quickly lyse when invaded by parasites and so the parasites die. In the case of the sickle mutation this occurs because of the sickling effect while with HbE and thalassemias it reflects the small and poorly hemoglobinized red blood cells.

There are many different hemoglobinopathies, but usually one type predominates in a given population; for example, black Africans will have HbS, South East Asians HbE and Mediterranean populations will have different thalassemias. Each protects against malaria but co-inheritance can cancel out this effect. Thus, HbS co-inherited with  $\alpha$  thalassemia removes the malaria protection because it makes the red blood cell abnormality less severe [34].

Genetic factors may also enhance the risk of infection. These are more subtle as they are thought to involve multiple genetic effects; i.e. QTLs (quantitative trait loci) that are difficult to detect. They have been sought by association (case control) studies and now by GWAS (genome wide association studies) (Chapters 2, 3). These studies have identified predisposition genetic loci to *N. meningitidis* meningitis, tuberculosis, HCV, leprosy and HBV. In the case of HBV it is the HLA locus that seems to be the key factor in predisposition and it is perhaps not coincidental that non-response following vaccination with HBV vaccine is more likely to occur in those with certain HLA types such as DRB1\*03 and DRB1\*07 HLA types [33].

### **Influenza**

The three RNA influenza viruses (A, B, C) are distinguished by their internal group-specific ribonucleoprotein. Only influenza A and B are medically significant, since epidemics or pandemics have not occurred with influenza



**FIGURE 6.6 Structure of the Influenza virus.** This RNA virus has two key surface glycoproteins: (1) Hemagglutinin (HA or H) – facilitates the entry of virus into host cells through attachment to sialic acid receptors, and (2) Neuraminidase (NA or N) – involved in the release of progeny virions from infected cells. The HA is the major determinant against which are directed neutralizing antibodies, and so also the target for influenza vaccines. In contrast, the NA is an important target for antiviral agents.

C. Influenza A has the potential to produce pandemics because it infects other species apart from humans, including birds, pigs and horses. Influenza B only infects humans and so its antigenic structure does not become sufficiently different to cause pandemics. In contrast, viruses such as measles undergo minimal antigenic variation with one infection giving life-long immunity.

The subtyping of the influenza A virus is based on its outer viral proteins, which include two important and distinct antigenic glycoproteins: *Hemagglutinin* (H – composed of 16 different types) and *neuraminidase* (N – nine different types) (Figure 6.6). Although the envelope antigens are capable of producing many different combinations (as seen in water birds), a smaller number are found in humans. To date only a few have been implicated in human to human spread (H1N1, H2N2, H3N2, H1N2, H5N1, H9N2 and H7N7) with highly pathogenic avian influenza subtypes found only in H5 and H7 subtypes (Figure 6.7) [35].

As the influenza A virus passes through its hosts, the most important of which in terms of

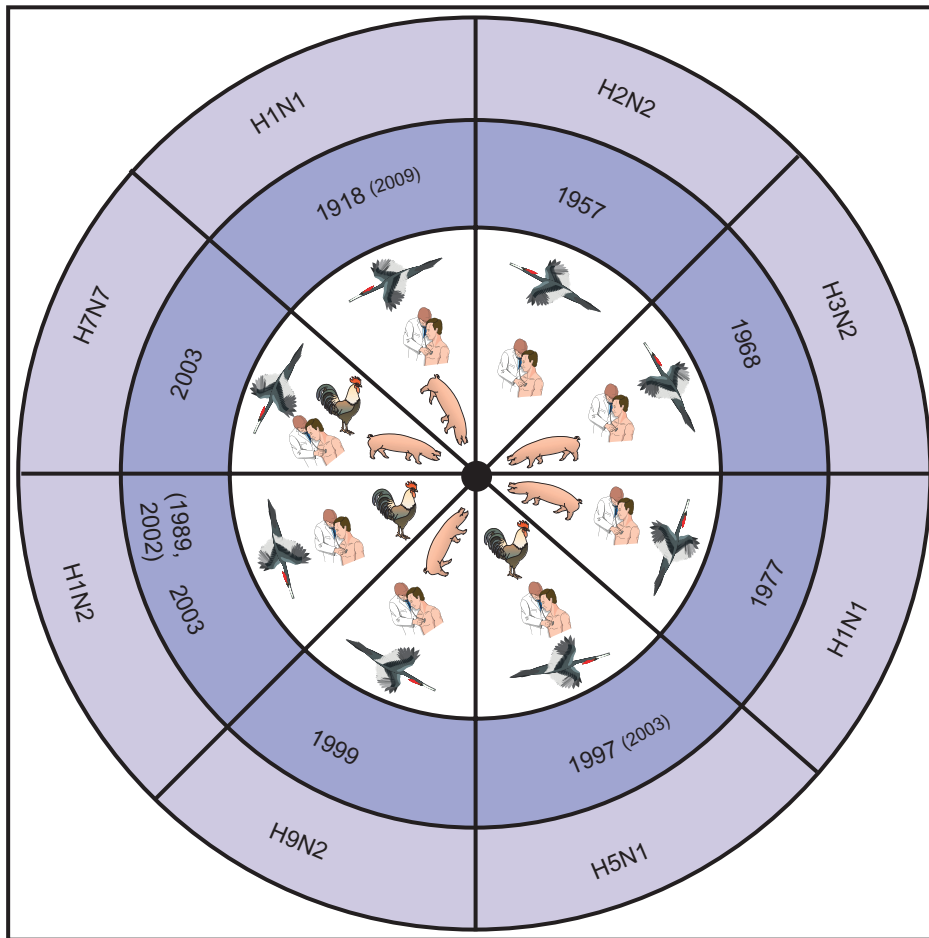
global spread are the water birds, it undergoes genetic changes. In the past 100 years there have been four influenza pandemics:

1. 1918 H1N1;
2. 1957 H2N2;
3. 1968 H3N2, and
4. 2009 H1N1.

A fifth outbreak (H5N1) has not been declared a pandemic but remains a concern.

*Avian influenza (avian flu, bird flu, H5N1, 1997 and re-emergence in 2003):* This remains a worldwide threat to health, with some regarding a H5N1 pandemic as being potentially more devastating than the 1918 Spanish flu outbreak. In 1997, the first cases of human infection from exposure to sick birds or their droppings were reported in Hong Kong, indicating that this virus subtype had jumped the species barrier. Eighteen patients were admitted to hospital and six died. Fortunately, the timely culling of over a million chickens controlled this particular outbreak. Today, H5N1 still causes outbreaks in chickens, and sporadic human infections continue to be reported, with a mortality of over 50%. In contrast to H1N1 swine flu and SARS (Severe Acute Respiratory Syndrome) that have been spread from human to human and through travel, the H5N1 bird flu remains relatively contained because spread is predominantly through chickens or other birds.

The common human influenza virus (H3N2) is highly contagious but rarely lethal. Avian flu in chickens (H5N1) is a particularly virulent type that can kill rapidly and causes widespread organ damage. Fortunately, it is not easily transmitted from birds to humans, and more importantly, human to human spread is poor. However, swapping genetic material, should an individual be co-infected with both, might produce a hybrid H5 (avian flu) N2 (human flu) virus with devastating effects. DNA sequencing of the viral genome from various outbreaks has shown that the virus continues to



**FIGURE 6.7** Major animal–human and human–human influenza outbreaks. Since the 1918 pandemic, a number of important outbreaks have been recorded (subtypes and dates are given as well as hosts involved). A worrying trend is the increasing numbers of new subtypes in humans, as well as an expanding animal involvement since 1997, in particular the domestic chicken.

mutate. This has implications for pathogenicity, as well as antiviral drug resistance, and having the right vaccine ready if needed. In this unpredictable environment, the value of rapid NAT diagnostics is crucial to detect early cases and for public health surveillance. The genes of the virus that caused the 1918 pandemic have been studied to better understand what makes an influenza virus virulent and capable of producing a pandemic [35].

*Spanish influenza (H1N1, 1918):* The virus from this pandemic, which killed about 40 million people, had not been isolated. Without a virus little research was possible, then the viral RNA sequence was determined using material from archival tissue, including formalin-fixed autopsy material. The sequence itself did not provide clues for why the Spanish influenza virus was so virulent, and so the next step was to reconstruct the viral coding segments and clone them

into plasmids. Individual genes from the H1N1 1918 virus were then introduced into a common laboratory viral strain and pathogenicity sought. Although the H and N glycoproteins were factors in the virulence of this virus, it was also shown that one of the RNA polymerase subunits known as PB1 was involved. Another subunit

(PB2) was then found to be important for viral transmissibility [35].

*Swine influenza (H1N1, later called H1N1(09), 2009):* After the appearance and then rapid disappearance of SARS (Box 6.4), followed by the concerns regarding the possibility of a H5N1 pandemic that did not occur (so far),

#### BOX 6.4

### SARS (SEVERE ACUTE RESPIRATORY SYNDROME).

This infection attracted a lot of publicity and provoked considerable fear when it emerged in China and then Hong Kong in 2003. SARS subsequently spread to many countries, producing around 700 deaths in the first half of 2003. This was at one time described as the first pandemic of the 21st century, but it never progressed beyond an epidemic because of effective public health measures effected by mid 2003 [36]. The social and economic impacts of this infection were considerable, including major disruptions to international travel. SARS was shown to be caused by a novel coronavirus (CoV) which was thought to have crossed the species barrier, although the animal reservoir for SARS took a while to find. It is now thought to be:

1. Masked palm civets – used for exotic food dishes in China, and
2. Horseshoe bats [37].

Traditional approaches such as viral culture, electron microscopy and serology helped to characterize the SARS virus. Nevertheless, SARS illustrated the value of NAT approaches in dealing with an emerging virus. Molecular testing enabled the following to be possible in a very short time frame:

1. Typing of the virus from two different countries (Taiwan and Hong Kong) showed that human to human spread had occurred;

2. Rapid whole genome sequencing of viral RNA enabled the development of PCR based diagnostic assays, and
3. In searching for animal reservoirs, RT-PCR based techniques were used. These allowed SARS-CoV to be detected, as well as identifying genetic differences between the human and animal virus.

The outbreak ended just as quickly as it started. Only occasional cases were reported in early 2004, and none after the end of April that year. However, there remain many unanswered questions including the inconsistent human to human transmission which might have been due to super-spreaders.

Another observation was the relatively large numbers of health workers who became infected. This became an issue when two of the nine persons infected in China in 2004 worked in a reference laboratory conducting research into the virus. A similar scenario was reported earlier in Singapore. The latter case was documented on RNA sequencing of the virus to be due to a contaminated laboratory culture that the scientist had been working with three days before showing signs of the infection. The WHO subsequently flagged the importance of laboratory containment when dealing with the SARS virus.



the world in 2009 was faced with another possible serious influenza outbreak. This outbreak was described as swine flu, because it was a well-recognized cause of influenza in pigs. The virus is related to the H1N1 virus that caused the Spanish flu, and can spread from person to person. The WHO declared a swine flu pandemic in June 2009. Vaccines were rapidly developed and stockpiles of antiviral drugs, particularly the two mentioned in Table 6.5, were released to the public. Rapid NATs requiring RT-PCR because it is an RNA virus were developed (see Table 3.3). This flu was a little unusual because it tended to be more severe in younger people, including children and pregnant women whereas deaths from seasonal flu involve mostly the elderly. Despite early concerns expressed by public health officials and considerable media hype, the WHO declared the H1N1 pandemic over in August 2010.

### Emerging and Re-Emerging Infections

Emerging (newly discovered, for example SARS – Box 6.4) and re-emerging (previously known, for example dengue virus) infections have increased significantly in the past 20 years. Many factors contribute including:

- Globalization, particularly increased travel and trade;
- Changes in human behavior, poverty and social inequality;
- Economic development, changes in the environment, weather and land use;
- Lapses in public health measures including those due to poverty or war;
- Complacency by communities or government;
- Mutations, selection and genetic reassortment in organisms;
- Bioterror.

Very few emerging infections represent novel pathogens. Most are re-emerging infections resulting from a change in the epidemiology or

virulence of a pathogen, or secondary to microbial adaptation. A review of the major infections in history provides some background to the emerging ones. They are:

1. Plague of Athens 430 BC;
2. Black death (*Y. pestis*) in 1340s;
3. French pox (syphilis) 1494;
4. Small pox 1520;
5. European cattle epidemics including anthrax, foot and mouth disease 1700s;
6. American plague (yellow fever) 1793;
7. Cholera pandemic in Paris 1832;
8. Measles outbreak in Fiji 1875;
9. Spanish influenza 1918, and
10. HIV-AIDS from 1981 [38].

### Zoonoses

Most emergent viruses are zoonotic – i.e. they are acquired from animals that are reservoirs of infection. This is particularly relevant in the modern world, where the consequences of easy migration, deforestation, agricultural practices, dam building and urbanization are making, and will continue to make, a major impact on the ecology of animals. For example, yellow fever is thought to have emerged in the New World as a result of the African slave trade which brought the mosquito *Aedes aegypti* in ships' water containers. More recently, *Aedes albopictus*, a potential vector for dengue virus, has become established in the USA following its conveyance from South East Asia in old car tires. With this, the threat of dengue in the North American continent has become real. Humans have populated rural areas to an increasing extent, as well as pursuing more outdoor recreational activities. There is also a growing trend for exotic animals to be kept as household pets. Changes in global climate may also contribute directly, through their effects on vegetation, insect and rodent populations.

Table 6.7 lists a number of zoonoses that have become established as new infectious diseases, or are emerging as problems for the future. Some

**TABLE 6.7** Some examples of zoonoses resulting in *new* human infections<sup>a</sup>.

Pathogen	(1) Clinical problems, (2) Emergence, (3) DNA applications
West Nile virus – RNA virus from <i>Flaviviridae</i> family (Genus <i>Flavivirus</i> ) – related to Yellow fever; Japanese encephalitis	(1) An asymptomatic febrile illness but can be complicated by meningitis, encephalitis or paralysis. Usually transmitted by mosquitoes. Also associated with blood or organ donation; pregnancy, lactation; infected needles or laboratory specimens. (2) Isolated in 1937 from Uganda and found in many parts of the world. Appeared in the USA in 1999, and has rapidly spread across North America. The virus is maintained by a bird-mosquito-bird cycle. (3) NAT is used to screen blood donors who may be asymptomatic carriers.
Monkeypox virus – DNA virus from <i>Poxviridae</i> family ( <i>Orthopoxvirus</i> ) – related to smallpox	(1) Self limited febrile illness with vesiculo-pustular eruptions. Confused with more serious illnesses and is spread animal to human or human to human. (2) Recognized in 1958 and remained localized to Africa until 2003 when it was detected in a mid-west USA outbreak. Traced back to rats imported from Africa to which native prairie dogs were exposed and became infected and then infected humans. Appears to be contained. Primary animal reservoir is the rat. (3) DNA characterization helped in identifying this virus as monkeypox.
Ebola virus (Ebola) – RNA virus from <i>Filoviridae</i> family ( <i>Ebolavirus</i> ) – related to Marburg virus	(1) Hemorrhagic fever in humans (mortality 50–90%). Example of increased human to animal contact in tropical forest with outbreaks generally resulting from the handling of infected dead animals. Humans highly contagious once disease established. (2) First isolated in 1976 from Sudan and Zaire. Since, sporadic outbreaks have occurred but remain in Africa. Animal host is unclear although bats are suspected as being natural reservoirs. (3) NAT assays for rapid and sensitive diagnostic tests described.
Lassa virus (Lassa fever) – RNA virus from <i>Arenaviridae</i> family ( <i>Arenavirus</i> )	(1) Hemorrhagic fever with 20% having severe multisystem disease. Virus excreted in human urine or semen for months post infection. (2) Endemic in west Africa since 1950s. Rodents are the primary reservoir and infect humans through fecal or urine contamination of food stores or if eaten. Human to human transmission occurs. (3) RT-PCR multiplex assay that can detect all important acute hemorrhagic fever viruses and provide information on viral loads has been described [39].
Hantavirus – RNA virus from <i>Bunyaviridae</i> family ( <i>Hantavirus</i> )	(1) Hemorrhagic fever with renal and pulmonary syndromes causing potentially fatal disorder. Infection occurs through exposure to aerosolized rodent excreta or bites. The aerosolization aspect makes this virus a particular concern for bioterror. (2) Isolated in 1979 in Korea. Now established within the Eurasian continent and the Americas. Outbreaks reported in the USA thought to be related to climatic changes, increasing vegetation and rodent population. (3) See Lassa fever.
Lyme disease – Bacterial spirochaete <i>Borrelia burgdorferi</i>	(1) Early non-specific malaise can be complicated by arthritis, neurologic and cardiac problems. Tick ( <i>Ixodes</i> spp.) transmitted disease. (2) First recognized in USA in 1957, since then reported in many countries. Mice, rodents and birds are the intermediate hosts. (3) DNA characterization useful for epidemiologic purposes, and to explain variable clinical features in different countries.
HIV – RNA virus from <i>Retroviridae</i> family ( <i>Lentivirus</i> )	(1) Serious acquired immunodeficiency disorder. (2) Cross species transmission from non-human primates followed by human to human spread. Detected in 1981. Evidence for the link between non-human primate and human disease includes: (i) Similar viral genomes; (ii) Prevalence in the natural host, and (iii) Geographic co-location. (3) NAT has been helpful in all phases of this particular disease from diagnosis to prognosis (in terms of viral load determination and detection of viral resistance).

(Continued)

TABLE 6.7 (Continued)

Pathogen	(1) Clinical problems, (2) Emergence, (3) DNA applications
Australian bat Lyssavirus (ABLV) – RNA virus from <i>Rhabdoviridae</i> family ( <i>Lyssavirus</i> ) Hendra virus – RNA virus from <i>Paramyxoviridae</i> family ( <i>Henipavirus</i> )	(1) Serious viral infections with high risk for fatal encephalitis first causing a problem in Australia in the mid 1990s. Lyssavirus closely related both serologically and molecularly to the rabies virus while Hendra is related to Nipah virus. (2) Bats are reservoirs for both viruses. Two ABLV deaths reported to date have resulted from a scratch or bite from an infected bat. Hendra infects humans via exposure to the bodily fluids of infected horses. (3) Both phenotypic (serology) and genotypic (PCR tests) available for ABLV and Hendra.
Chikungunya (CHIKV) – RNA virus from <i>Togaviridae</i> family ( <i>Alphavirus</i> )	(1) Can result in severe illness comparable to dengue fever followed by arthralgias that can last for years. Transmitted to humans by <i>Aedes</i> mosquito bites. (2) Endemic to tropical Africa (first isolated in Tanzania in 1953) and Asia although recently outbreaks seen in Western Pacific, Europe and India. Main reservoirs are monkeys. (3) Traditional tests take time or may give false positives (serology). RT-PCR useful and gives rapid result.

<sup>a</sup>BSE, CJD, and vCJD are dealt with in the text.

of these are newly acquired in the west, while others remain endemic to specific countries. However, any disease may be spread through international travel, or the mass dislocation of large populations through civil unrest. There is also an increasing possibility that a number of pathogens could be used for bioterrorism. Some of the zoonoses associated with a viral hemorrhagic clinical picture can be confused with other clinical infections including malaria, leptospirosis, and *N. meningitidis* and in these potentially fatal conditions, a rapid screening test is essential. In terms of bioterrorism and the differential diagnosis of hemorrhagic fevers, NAT assays are presently the only option to allow rapid and sensitive diagnostic tests to be developed. If new therapeutics are required, the first step will be nucleic acid sequence analysis of the microorganisms' genomes so that it can be classified and identified. Next, potential targets for vaccines or drug therapies can be established.

## GLOBAL HEALTH

In an era of personalized medicine, one should not lose sight of how molecular medicine

can be used to improve global health. Cheaper drugs and vaccines for all communities is an important benefit that should come from molecular-based technologies. Another would be better NATs. In this respect it is intriguing to recall how direct-to-consumer DNA testing (Chapter 5) makes effective use of the Internet. Could the Internet be one way to improve accessibility for disadvantaged communities or those in rural and remote regions? Consideration of how genomics can play a part in the bioeconomy, with its potential to generate income, improve food production and sustain a better environment, are some of the challenges now being taken up by bodies such as the OECD.

## Non-Communicable Diseases

A large part of this chapter has dealt with infectious diseases and how these impact on individuals, communities and ultimately global health. To complete the story, it is necessary to consider non-communicable diseases since, apart from their primary effect on health and well being, they can also contribute to a communities' vulnerability to infectious diseases (Table 6.8).

**TABLE 6.8** Some global health challenges [21,40–41].

Communicable diseases	Non-communicable diseases
Major infectious disorders – AIDS, TB and malaria	Hypertension
Emerging (new) pathogens, e.g. Hendra, Australian bat lyssavirus	Tobacco and alcohol
Re-emerging (old) pathogens, e.g. <i>C. difficile</i> , mumps	Hyperglycemia, physical inactivity, overweight and obesity
Influenza A (H1N1 and H5N1 strains) and resistance	Childhood underweight
Zoonoses accounting for about 60% of emerging and re-emerging infections	Unsafe water, poor sanitation and hygiene
Antibiotic resistance, e.g. multidrug resistance, extensively resistant TB, vancomycin resistant enterococcus, and drug resistance, e.g. malaria	Indoor smoke from solid fuels (low-middle income countries) or urban outdoor air pollution (high income countries)
Agents that could be used for bioterror, e.g. anthrax, plague, smallpox	Suboptimal breast feeding, low fruit or vegetable intake

A *Perspective* on global non-communicable diseases makes some sobering observations including:

1. 60% of all deaths are due to chronic diseases, with most occurring in low to middle income countries with a disproportionate number of young people dying during their productive years;
2. Non-communicable diseases are likely to have a more detrimental effect on global economic development than fiscal crises, natural disasters or pandemic influenza;
3. In the next 10 years, it is projected that China (as one example) will lose \$558 billion in national income because of preventable heart disease, stroke and diabetes, and

4. To address these problems it is essential to have better evidence-based decision making, more effective regulation and behavioral interventions that are known to work.

The need to shift focus more to community-based prevention and concentrate less on attempting to cure a problem once it is established has already been highlighted [40].

## Obesity

A number of the non-communicable health problems listed in Table 6.8 have obesity as a contributing factor. In the USA obesity continues to be a major health challenge; 2003–2004 estimates indicated that 66% of the US population was overweight, and 32% obese, as defined by a BMI  $\geq 30$  kg/m<sup>2</sup>. Another estimate is that 50% of the adults in the USA will be clinically obese by 2030 [42,43].

Current understanding is that most cases of obesity are caused by a mix of genetic and environmental factors, although their relative contributions remain to be determined. The rapid development of obesity worldwide can only be an environmental effect. Nevertheless, many people in the same environment have not developed obesity and so genes must play a role. Comparisons between monozygotic and dizygotic twins, as well as other studies, show greater concordance for the BMI (a surrogate measure for obesity), i.e. there is an important genetic component to obesity, with estimates indicating that this is a strong effect (around 80%) [43].

One hypothesis, which has been around for 50 years, captures both genes and environment. It suggests that genes important for metabolism in humans evolved over time to respond to periods of famine. These so called *thrifty genes* allowed hunter-gatherer populations to process food into fat deposits during times of plenty, so that they could survive when food was not available. Today, these same genes

respond inappropriately when food is readily available all year round, and so obesity results. Evidence for this genetic evolutionary effect is still awaited. Other hypotheses include:

1. Fetal programming (perhaps via epigenetic changes) with maternal nutrition a key factor in how the child will grow postnatally;
2. Sedentary lifestyle, i.e. diet and lifestyle are the main contributors and from the genetics perspective this would put the focus onto metabolic enzymes;
3. Increased reproductive fitness, since the number of offspring is positively correlated with the BMI of women – i.e. adiposity increases fertility, and
4. Many others [43].

The public health response to the obesity epidemic is focused on eating less, avoiding fast foods and exercising more. However, this approach is not working. Can a more personalized genomics strategy help? Will a scientifically plausible understanding of how diet, the environment and obesity interact allow governments and individuals to take a more effective approach? One way to pursue this would be to know more about the genes involved in obesity.

### **The Genetics of Obesity**

Our current understanding of genes and obesity is still rudimentary, so medical or motivational interventions cannot be tested. At the genetic level, obesity can be considered in three groups:

1. Monogenic, Mendelian defects, such as mutations in the melanocortin-4 receptor gene (*MC4R*) leading to an autosomal dominant cause for obesity in up to 6% of individuals, particularly those with more severe forms and earlier ages of onset (Box 6.5);
2. Syndromal disorders such as Prader-Willi syndrome, Bardet-Biedl syndrome and Pseudohypoparathyroidism type 1A, and

3. Complex but common forms of obesity for which the traditional association or GWAS have been used to identify risk alleles [44].

Genes or gene loci implicated in obesity have been listed in a Human Obesity Gene Map last updated in 2005 [45]. This map provides a summary of published data that are not necessarily confirmed or authenticated but gives a flavor of the rich genetic heterogeneity expected with a complex phenotype such as obesity. Observations made about the 2005 human obesity map include:

1. 176 cases involving obesity in humans are due to single gene mutations in 11 genes;
2. 253 genetic loci have been reported for obesity from genome wide scans;
3. There are 426 findings of positive associations with 127 candidate genes;
4. Association studies in 22 genes have been replicated at least five times, and
5. There are putative obesity loci on all chromosomes except Y.

### **Microbiome and Obesity**

It is intriguing to recall the observation in Chapter 4 that the gut metagenome shows a characteristic alteration in obese subjects, and so the microbial flora may play a role in obesity that is independent of net calorie intake. In obese humans and animals (mouse, rat and pig) the ratio of the two major bacterial divisions in the gut shows a predominance of *Firmicutes* over the *Bacteroides*. This is likely to be a primary rather than secondary effect, because when germ-free mice were fed the microbiota derived from lean or obese mice, the phenotype of the recipient mice moved towards that of the donor mouse – i.e. the obese or lean phenotype was transmissible via the microbiome. One mechanism for this observation may be that the obese microbiome can extract more energy from food [46]. New targets for interventions may be found as the metagenomics story unfolds and more is found about the gut

## BOX 6.5

## GENES AND OBESITY.

Apart from the *MC4R* example given, other genes associated with obesity have a recessive mode of inheritance. They include mutations causing deficiency in leptin and its receptor (*LEP*, *LEPR*) which act via the hypothalamus to control appetite and energy expenditure. One report, concerning a child with congenital leptin deficiency, described how a sustained reduction in weight occurred following treatment with recombinant human leptin. Other genes in the leptin-melanocortin pathway are also implicated including *POMC* and *PCSK1*. A human gene *FTO* was shown to be implicated strongly with the BMI (body mass index) in a genome wide association study involving subjects with type II diabetes. This has been replicated in other studies and appears to be reflecting common SNP polymorphisms in intron 1, with the risk allele highly prevalent in the general population. European carriers who are homozygous for the risk allele weigh on average 3kg more. Some clues to *FTO* gene function include:

1. *Fto* null mice are protected from obesity by increased energy expenditure;
2. *FTO* expression in humans is highest in the brain, particularly the cerebral cortex, and
3. Duplication of a chromosomal region containing *FTO* (and other genes) was associated with mild obesity and mental retardation in a case study.

It was reported recently that a reduction in brain volume in healthy elderly individuals was also associated with the same *FTO* allele for obesity. Perhaps this is not surprising since obesity is also a risk factor in cognitive decline and dementia. Very rare monogenic causes of obesity include mutations in genes associated with hypothalamic function such as *SIM1*, *BDNF* and *NTRK2*. These may lead to abnormalities in energy balance resulting in hyperphagia and a net positive energy intake [43,44].

flora and its effects on a range of issues including obesity and inflammation.

### Nutrigenetics and Nutrigenomics

Nutrition is a key environmental variable and so any starting point in understanding obesity must encompass nutrition, including its various genetic components. There is a parallel here with pharmacogenetics. Conventional dietary guidelines take consideration of age, sex, height, weight and level of physical activity but not genetic variability. Many of these parameters are used to determine drug dosage, although it is now clear that genetic variability

also plays an important role (Chapter 3). A more personalized approach becomes possible through *nutrigenetics* – how individuals respond differently (because of genetic variation) to the same diet, for example, through changes in blood pressure or serum cholesterol, and *nutrigenomics* – the role of nutrients and bioactive food compounds in gene expression. The ultimate goal is the development of personalized nutrition options to ensure health and prevent disease [47]. Overarching these goals is the incredible diversity of genetic, cultural and environmental considerations in diet. Nutrigenomics can be approached through many of the omics including genomics,

epigenomics, transcriptomics, proteomics, metabolomics and so on.

### **Diet and Cancer**

One can be sure of controversy and robust debate when the influence of diet, nutraceuticals (nutrition + pharmaceutical), complementary medicines or food additives are discussed in relation to cancer development. Knowledge of the link between cancer and diet is not new and numerous research studies provide conflicting data. This is not surprising since individual genetic variability will make the small, multiple but cumulative effects of diet on DNA damage difficult to measure or even replicate, just as association-based studies looking for genetic factors in complex diseases produce conflicting results.

One example is vitamin D deficiency, which is said to cause cancer, although this is very controversial. The US National Cancer Institute confirms a knowledge gap here, stating it does

not recommend *for or against* the use of vitamin D supplements in reducing the risk of cancer. The D2 and D3 forms of vitamin D need to be metabolized to the active 1,25-dihydroxy-vitamin D and this involves a number of enzymes (including cytochrome P450 discussed earlier in relation to drug metabolism in Chapter 3). The role of vitamin D in cancer may be better understood through a molecular approach. This is important in view of the successful public health campaigns in reducing the risk of sun-related skin cancers. Interventions recommended include the generous application of sunscreens, avoidance of sun and the wearing of wide brimmed hats, particularly in children. While successful in preventing skin cancers, there is concern (although this is controversial) that vitamin D deficiency may result. If so, there are risks to consider in terms of rickets and related bone problems, and potentially cancer.

The nutrition of cancer cells is also an area of interest. A relevant observation is known as

**TABLE 6.9** Delivering growth and labor productivity through genomics [49].

Activity	Examples
Agriculture	Conventional agronomic practices have helped to increase global food yields but more is needed as the world's population increases. Genetic-based knowledge is now being added to overcome roadblocks in productivity. A major step forward occurred when whole genome DNA sequences of many plants and staple foods such as rice were published. Salinity, drought and uncontrolled flooding are some of the challenges for rice growers. Whole genome sequences are now being interrogated to identify genes that might overcome these problems without necessarily going the full but controversial next step which are GM (genetically modified) crops.
Livestock	As living standards improve so does the expectation that more protein in the form of meat will become available as food. Like agriculture, the traditional animal breeding approach has led to better yields except for fish. Mapping in the 1990s to identify genes that would enhance breeding was also effective with the more powerful SNP mapping becoming an improvement on this in the 2000s. Today, whole genome sequencing has been completed in the pig, chicken and cattle and is expected to identify important genes to improve breeding and meat yields.
Alternative fuels	Solar and wind power are being used as alternative energy sources although air transport still relies on petroleum fuels. There is now considerable interest in identifying genes in the cow rumen or the termite gut to find new enzymes that can be used to digest wood isolated from various crops and so produce sugars that can be fermented into ethanol for fuel. Algae may also be induced to overexpress ethanol-producing genes and for this all that is needed is sunlight. As a bonus algae will take in carbon dioxide from the atmosphere.

the *Warburg effect*. O. Warburg was awarded the 1931 Nobel Prize in Physiology or Medicine for discovery of cytochrome C oxidase. He also showed that cancer cells produce lactic acid from glucose even under non-hypoxic conditions; an observation that now bears his name. This is considered to reflect abnormal regulation of glycolysis, since this pathway is very active compared to normal cells, even in the presence of sufficient oxygen [48]. This finding might have implications for new cancer therapy targets and help us to understand better how genes are involved in cancer causation.

## Bioeconomy

The OECD broadly defines bioeconomy as “*the set of economic activities relating to the invention, development, production and use of biological products and processes*”. It makes the prediction that biotechnology (in primary production, health and industry) can offer solutions that will lead to the emergence of a bioeconomy. The OECD as an economy-based organization considers greater social benefits globally will come from improving sustainable growth without depleting resources, and labor productivity. The latter can be enhanced through innovation, which is particularly suited to genomics as many of the future developments will be delivered *in silico* (Chapter 4) and so expensive infrastructure is not necessary. Some examples of how the bioeconomy will benefit from genomics and other omics can be found in Table 6.9.

The expectation is that the bioeconomy can be used to make substantial socioeconomic contributions to OECD and non-OECD countries, and from this will come better health outcomes, improved productivity of agriculture and industrial processes and enhanced environmental sustainability. In an attempt to optimize the potential of the bioeconomy, the OECD has published a long term (2030) policy agenda [50].

## References

- [1] CDC website on public health genomics. <http://198.246.98.21/genomics/about/AAG/index.htm>
- [2] Starfield B, Hyde J, Gervas J, Heath I. The concept of prevention: a good idea gone astray? *Journal of Epidemiology and Community Health* 2008;62:580–3.
- [3] WHO screening criteria as interpreted in one jurisdiction (Australia). [www.health.gov.au/internet/screening/publishing.nsf/Content/pop-based-screening-fwork/\\$File/screening-framework.pdf](http://www.health.gov.au/internet/screening/publishing.nsf/Content/pop-based-screening-fwork/$File/screening-framework.pdf)
- [4] Parsons EP, Bradley DM. Newborn screening programmes. In: *Encyclopedia of Life Sciences (ELS)*. Chichester: John Wiley & Sons, Ltd.; 2008.
- [5] Khoury MJ, McCabe LL, McCabe ERB. Population screening in the age of genomic medicine. *New England Journal of Medicine* 2003;348:50–8.
- [6] Castellani C, Macek M, Cassiman J-J, et al. Benchmarks for cystic fibrosis carrier screening: a European consensus document. *Journal of Cystic Fibrosis* 2010;9:165–78.
- [7] Bonham VL, Dover GJ, Brody LC. Screening student athletes for sickle cell trait – a social and clinical experiment. *New England Journal of Medicine* 2010;363:997–9.
- [8] Wilcken B, Wiley V. Newborn screening. *Pathology* 2008;40:104–15.
- [9] McBride CM, Koehly LM, Sanderson SC, Kaphingst KA. The behavioral response to personalized genetic information: will genetic risk profiles motivate individuals and families to choose more healthful behaviors? *Annual Review of Public Health* 2010;31:89–103.
- [10] UK 2009 NICE guidelines in familial hypercholesterolaemia. [www.nice.org.uk/nicedia/pdf/CG071NICEGuideline.pdf](http://www.nice.org.uk/nicedia/pdf/CG071NICEGuideline.pdf)
- [11] Heart UK; familial hypercholesterolemia git Toolkit to implement NICE guidelines. [www.heartuk.org.uk/FHToolkit/](http://www.heartuk.org.uk/FHToolkit/)
- [12] Genetics in the workplace: implications for occupational safety and health Nov 2009. Department of Health and Human Services. Centers for Disease Control and Prevention. National Institute for Occupational Safety and Health. <http://origin.cdc.gov/niosh/docs/2010-101/pdfs/2010-101.pdf>
- [13] McCanlies EC, Kreiss K, Andrew M, Weston A. HLA-DPB1 and chronic beryllium disease: A HuGE review. *American Journal of Epidemiology* 2003;157:388–98.
- [14] McKee AC, Cantu RC, Nowinski CJ, et al. Chronic traumatic encephalopathy in athletes: progressive tauopathy after repetitive head injury. *Journal of Neuropathology and Experimental Neurology* 2009;68:709–35.



- [15] Pleasance ED, Stephens PJ, O'Meara S, et al. A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* 2010;463:184–90.
- [16] Relman DA. Microbial genomics and infectious diseases. *New England Journal of Medicine* 2011;365:347–57.
- [17] Lapiere P, Gogarten JP. Estimating the size of the bacterial pan-genome. *Trends in Genetics* 2009;25:107–10.
- [18] Williams CH, Stanway G. Viruses: genomes and genomics. In: *Encyclopedia of Life Sciences (ELS)*. Chichester: John Wiley & Sons, Ltd.; 2009.
- [19] Belyi VA, Levine AJ, Skalka AM. Unexpected inheritance: multiple integration of ancient Bornavirus and Ebolavirus / Marburgvirus sequences in vertebrate genomes. *PLoS Pathogens* 2010;6:e1001030.
- [20] Peleg AY, Hooper DC. Hospital-acquired infections due to gram-negative bacteria. *New England Journal of Medicine* 2010;362:1804–13.
- [21] Pang T. Germs, genomics and global public health. How can advances in genomic sciences be integrated into public health in the developing world to deal with infectious diseases. *HUGO Journal* 2009;3:5–9.
- [22] Lee JH, Jeong SH, Cha S-S, Lee SH. New disturbing trend in anti-microbial resistance of gram-negative pathogens. *PLoS Pathogens* 2009;5:e1000221.
- [23] Kumarasamy KK, Toleman MA, Walsh TR, et al. Emergence of a new antibiotic resistance mechanism in India, Pakistan, and the UK: a molecular, biological and epidemiological study. *The Lancet Infectious Diseases* 2010;10:597–602.
- [24] Travassos MA, Laufer MK. Resistance to antimalarial drugs: molecular, pharmacologic and clinical considerations. *Pediatric Research* 2009;65:64R–70R.
- [25] Kumar S, Kumar A, Dixit VK. Direct detection and analysis of vacA genotypes and cagA gene of *Helicobacter pylori* from gastric biopsies by a novel multiplex polymerase chain reaction assay. *Diagnostic Microbiology and Infectious Disease* 2008;62:366–73.
- [26] Alcaide F, Coll P. Advances in rapid diagnosis of tuberculosis disease and anti-tuberculous drug resistance. *Enfermedades Infecciosas y Microbiología Clínica* 2011;29(Supl 1):34–40.
- [27] WHO 2010 recommendations: Screening donated blood for transfusion-transmissible infections. <http://www.who.int/bloodsafety/ScreeningDonatedBloodforTransfusion.pdf>
- [28] Aguzzi A, Calella AM. Prions: protein aggregation and infectious diseases. *Physiological Reviews* 2009;89:1105–52.
- [29] Oldani A, Cormont M, Hofman V, et al. *Helicobacter pylori* counteracts the apoptotic action of its VacA toxin by injecting the CagA protein into gastric epithelial cells. *PLoS Pathogens* 2009;5:e1000603.
- [30] Rasko DA, Sperandio V. Anti-virulence strategies to combat bacteria mediated disease. *Nature Reviews Drug Discovery* 2010;9:117–28.
- [31] Simjee S, editor. *Foodborne diseases*. New Jersey: Humana Press; 2007.
- [32] Clements ACA, Soares Magalhaes RJ, Tatem AJ, Paterson DL, Riley TV. *Clostridium difficile* PCR ribotype 027: assessing the risk of further worldwide spread. *The Lancet Infectious Diseases* 2010;10:395–404.
- [33] Kaslow RA, Shrestha S, Tang JJ. Susceptibility to human infectious diseases, genetics of. In: *Encyclopedia of Life Sciences (ELS)*. Chichester: John Wiley & Sons, Ltd.; 2008.
- [34] Penman BS, Pybus OG, Weatherall DJ, Gupta S. Epistatic interactions between genetic disorders of hemoglobin can explain why the sickle-cell gene is uncommon in the Mediterranean. *Proceedings of the National Academy of Sciences of the USA* 2009;106:21242–21246.
- [35] Tumpey TM, Belser JA. Resurrected pandemic influenza viruses. *Annual Reviews of Microbiology* 2009;63:79–98.
- [36] WHO SARS risk assessment and preparedness framework October 2004. [http://www.who.int/csr/resources/publications/CDS\\_CSR\\_ARO\\_2004\\_2.pdf](http://www.who.int/csr/resources/publications/CDS_CSR_ARO_2004_2.pdf)
- [37] Shi Z, Hu Z. A review of studies on animal reservoirs of the SARS coronavirus. *Virus Research* 2008;133:74–87.
- [38] Morens DM, Folkers GK, Fauci AS. Emerging infections: a perpetual challenge. *The Lancet Infectious Diseases* 2008;8:710–9.
- [39] Trombley AR, Wachter L, Garrison J, et al. Comprehensive panel of real-time TaqMan™ polymerase chain reaction assays for detection and absolute quantification of Filoviruses, Arenaviruses and New World Hantaviruses. *American Journal of Tropical Medicine and Hygiene* 2010;82:954–60.
- [40] Narayan KMV, Ali MK, Koplan JP. Global noncommunicable diseases – where worlds meet. *New England Journal of Medicine* 2010;363:1196–8.
- [41] USA's NIAID summary of emerging and re-emerging infections. [www.niaid.nih.gov/topics/emerging/pages/list.aspx](http://www.niaid.nih.gov/topics/emerging/pages/list.aspx)
- [42] Agurs-Collins T, Khoury MJ, Simon-Morton D, Olster DH, Harris JR, Milner JA. Public health genomics: translating obesity genomics research into population health benefits. *Obesity* 2008;16(S3):S85–94.
- [43] Walley AJ, Asher JE, Froguel P. The genetic contribution to non-syndromic human obesity. *Nature Reviews Genetics* 2009;10:431–42.
- [44] Ho AJ, Stein JL, Hua X, et al. A commonly carried allele of the obesity-related *FTO* gene is associated

- with reduced brain volume in the healthy elderly. *Proceedings of the National Academy of Sciences of the USA* 2010;107:8404–9.
- [45] Rankinen T, Zuberi A, Chagnon YC, et al. The Human Obesity Gene Map: the 2005 update. *Obesity* 2006;14:529–644.
- [46] Ley RE. Obesity and the human microbiome. *Current Opinion in Gastroenterology* 2010;26:5–11.
- [47] Fenech M, El-Sohehy A, Cahill L, et al. Nutrigenetics and nutrigenomics: Viewpoints on the current status and applications in nutrition research and practice. *Journal of Nutrigenetics and Nutrigenomics* 2011;4:69–89.
- [48] Koppenol WH, Bounds PL, Dang CV. Otto Warburg's contributions to current concepts of cancer metabolism. *Nature Reviews Cancer* 2011;11:325–37.
- [49] OECD's / HUGO's Symposium on genomics and bioeconomy, Montpellier France 17 May 2010. [www.oecd.org/document/41/0,3343,en\\_2649\\_34537\\_45430633\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/41/0,3343,en_2649_34537_45430633_1_1_1_1,00.html)
- [50] OECD's: The bioeconomy to 2030: designing a policy agenda. [http://www.oecd.org/document/56/0,3746,en\\_2649\\_36831301\\_36960312\\_1\\_1\\_1\\_1,00.htm](http://www.oecd.org/document/56/0,3746,en_2649_36831301_36960312_1_1_1_1,00.htm)

Note: All web-based references accessed on 21 Feb 2012.