Past and Present: How have infections shaped the human genome?

Infections caused by pathogenic viruses and bacteria have significantly shaped the human genome. The key process by which this is understood to have happened is natural selection. In this essay, I will address the difficulties involved in identifying signatures of natural selection, the factors that affect evolution of the genome and finally give some examples of where the human genome has been modified by pathogens.

Key terms and definitions

Scientists, such as Pasteur and Koch, paved the way for the study of microbes and how they cause disease. However, it was Haldane in 1949 who suggested that microorganisms influenced the distribution of human traits across the world and made the link between infections and natural selection.[1] Following on from this, links have now been made between genetic mutations at a population level and different immune responses to pathogens.[2]

According to Brinkworth, evolution of the human immune system has been described as 'pathogen-mediated selection on the genome that affects immune function.'[3] Pathogen-mediated selection describes the evolution of an organism's genome due to infection. Pathogens exert 'selective pressure' on the human genome. Selective pressure is any cause that can increase or decrease the reproductive ability of an organism which encourages evolution of the organism over time. As a result of selective pressure from pathogens, different alleles are favoured that are advantageous for fighting a particular disease. Of course, pathogens have a short reproductive cycle, so they are able to adapt faster to new environments and conditions. They are themselves under even stronger selective pressure than the organism they are infecting,[4] which in turn influences their effect on the human genome.

Through natural selection, mutations of alleles that are advantageous for survival are favoured, and therefore organisms with those mutations are more likely to survive and reproduce. There are two main types of natural selection: directional selection and balancing selection. Directional selection can be positive or negative. In positive directional selection, advantageous alleles are favoured and are increased in frequency. This can be identified through observation of high frequencies of new allele variations within a population. Negative directional selection describes decreasing frequencies of unhelpful or disadvantageous alleles. Balancing selection is where diversity is favoured, and there are multiple alleles at a locus.[5]

Difficulties with identifying signatures of selection

Signatures of natural selection are 'unusual patterns of allele frequencies that mark a selected locus.'[6] There are quite a few difficulties with identifying signatures of selection in the human genome related to infections. One reason for this is because it is hard to know whether every gene affected is specifically an 'immune gene' i.e., one involved in the immune response.

Another reason that it is hard to find specific signals of genome changes due to pathogens is because of a concept called immune redundancy. This is the idea that there are multiple receptors with slightly different DNA coding that overlap with each other and recognise the same pathogens.

---

[1] Barreiro, L. and Quintana-Murci, L. (2010) 'From evolutionary genetics to human immunology: how selection shapes host defence genes', *Nature Reviews Genetics,* 11, pp 17-30. Available at: From evolutionary genetics to human immunology: how selection shapes host defence genes | Nature Reviews Genetics

[2] Brinkworth, J F. (2017) 'Infectious Disease and the Diversification of the Human Genome', *Human Biology,* 89(1), pp 47-65. Available at: Infectious Disease and the Diversification of the Human Genome on JSTOR

[3] Brinkworth, J F (2017) 'Infectious Disease'

[4] Brinkworth, J F (2017) 'Infectious Disease'

[5] Karlsson, EK., Kwiatkowski, DP., and Sabeti, PC. (2014) 'Natural selection and infectious disease in human populations' *Nature Reviews Genetics,* 15, pp 379-393. Available at: Natural selection and infectious disease in human populations | Nature Reviews Genetics

[6] Karlsson, EK., Kwiatkowski, DP., and Sabeti, PC. (2014) 'Natural selection and infectious disease'

This is beneficial to humans as it means that there are several ways for the immune system to fight off infection, but it does complicate the search for finding specific evolutionary signals.[7]

A final difficulty is a factor called 'immune gene pleiotropy.' This is the idea that immune genes are not only involved in immune responses but also in other systems such as the reproductive and nervous systems. Therefore, it is hard to know whether a finding of high incidence is due to selective pressure by a pathogen or due to the gene's other functions.[8]

Factors affecting pathogen-mediated selection

There are a variety of factors that influence how pathogens have affected the human genome. There are differences in exposure to different pathogens due to human migration to various parts of the world, the lifestyles different people lead, and the exposure to different animals and pathogens that have been introduced to the human population. These all affect how pathogen-mediated selection may occur differently in different populations.[9]

Furthermore, signatures of selection are affected by the age of the pathogen, where (and how many places) in the world it is found, and how severe the disease can be. For example, *Yersinia pestis*, which causes the plague, has been demonstrated to have genomic impact. Studies show that alleles which allow a level of plague resistance were found in the highest frequency in areas most affected by the plague and therefore are examples of 'plague-mediated selection.'[10]

Genome shaping is also primarily led by natural selection from host-pathogen interactions. This is slowed in species with long lifetimes, slow reproductive cycles, and complex immune systems, especially humans.

A further factor to consider is that different pathogens might have influenced the same gene. This means that it can be difficult to pinpoint the exact effects of a single pathogen as there may have been more than one change at a single point.[11]

Finally, our immune system function is influenced by the presence of resident microorganisms. Take, for example, the production of lymphocytes, which are white blood cells important in the immune response against infection. Resident microorganisms are present in lymphoid tissue. They are important when producing emergency neutrophils which are white blood cells that kill and digest pathogens. Microorganisms create an environment that is competitive and so make it hard for pathogens to establish themselves as a threat in the body.[12]

How the genome has been shaped by pathogens

There is a plethora of ways that the human genome has been shaped by infectious diseases as outlined by Brinkworth:[13]

- Mutations in sections of DNA that code for proteins (i.e. coding regions) result in changes in the proteins produced e.g. mutations in the coding of receptors on plasma membranes mean that pathogens may not be able enter cells.
- Genomic structure can change for example the shortening of telomeres – the ends of chromosomes.
- Different populations have different immune responses to pathogens.
- Genomes can take nucleic acids of pathogens into themselves and incorporate them into the DNA, for example when viruses insert their genetic information into host cells.

[7] Brinkworth, J F (2017) 'Infectious Disease'
[8] Brinkworth, J F (2017) 'Infectious Disease'
[9] Brinkworth, J F (2017) 'Infectious Disease'
[10] Brinkworth, J F (2017) 'Infectious Disease'
[11] Barreiro, L. and Quintana-Murci, L. (2010) 'From evolutionary genetics'
[12] Brinkworth, J F (2017) 'Infectious Disease'
[13] Brinkworth, J F (2017) 'Infectious Disease'

Genes that are directly involved in immunity are the most obvious targets for natural selection. Natural selection has been found to target genes that play a part in both adaptive and innate immunity.[14] Innate immunity, also known as natural immunity, is part of the non-specific, primary immune response that humans are born with. Adaptive immunity, also known as acquired immunity, is the immunity gained against a certain pathogen because of clonal selection and expansion.[15] Clonal selection is the process of finding the lymphocytes with the correct antigens to neutralise a pathogen, and clonal expansion is the mass production of these lymphocytes to fight the infection.

However, some genetic mutations leading to disease resistance are not genes that are directly involved in dealing with immunity. For example, the removal of alpha-globin genes in red blood cells which leads to thalassemias (conditions affecting red blood cells) and gives immunity to malaria.[16] Sickle cell disease and sickle cell traits (SCTs) have been found to be more prevalent in countries in Africa, where malaria is a common disease. It was discovered that those with the sickle cell trait have more immunity to malaria than those without. It was suggested that where malaria is a more common disease, the number of thalassemias increased as they gave people a natural immunity to malaria.

Malaria is a classic example of pathogen-mediated selection of the human genome. For example, genes involved with erythrocyte function such as G6PD and HBB genes and immune response genes such as HLA genes were affected. HBB genes can have three amino acid changes that give different amounts of protection to malaria. One of the most common ones being the HbS allele mutation. Although this results in sickle cell disease, the individual has a ten times lower chance of developing severe malaria.[17]

At the HBB sickle cell locus, individuals who are heterozygous carriers, with one malaria-positive allele and one malaria-negative allele, are both malaria resistant and do not have sickle cell disease. This is an example of balancing selection. An example of positive selection is a mutation in the Duffy antigen gene (DARC). The Duffy antigen gene encodes receptors used by malaria parasites to enter red blood cells.[18]The mutation, with increased incidence in Africa, makes it more difficult for the malaria pathogens to enter host cells, and therefore the severity of the disease is decreased. Barreiro and Quintana-Murci researched the differences between humans and primates in their susceptibility to disease, and their immune responses. They compared the protein sequences between the species and were able to identify different proteins that evolved quickly. This meant they could see that proteins with functions linked to immunity were 'targets of positive selection.'[19]

Another example of pathogen-mediated changes in the human genome is found in the human population in the Ganges River Delta in response to cholera pathogens. It was found that those with blood type O were at a higher risk of contracting it more severely or dying. Their population now has the lowest rate of blood type O in the world.[20] This is an example of negative selection where individuals with the disadvantageous allele had lower reproductive ability, and therefore that allele decreased in frequency in the population.

The blood type antigens are most commonly found on the surface of erythrocytes, however they are also found on the surface of other cells, including the epithelial cells lining the intestinal wall Researchers from Washington University found that in people with blood type O, the cholera

---

[14] Cagliani, R. and Sirona, M. (2013) 'Pathogen-Driven Selection in the Human Genome', *International Journal of Evolutionary Biology,* 2013. Available at: Pathogen-Driven Selection in the Human Genome (hindawi.com)

[15] Gleichmann N. (2020) 'Innate vs Adaptive Immunity' *Technology Networks, Immunology and Microbiology.* Available at: Innate vs Adaptive Immunity | Technology Networks

[16] Brinkworth, J F (2017) 'Infectious Disease'

[17] Barreiro, L. and Quintana-Murci, L. (2010) 'From evolutionary genetics'

[18] Karlsson, EK., Kwiatkowski, DP., and Sabeti, PC. (2014) 'Natural selection and infectious disease'

[19] Barreiro, L. and Quintana-Murci, L. (2010) 'From evolutionary genetics'

[20] Karlsson, EK., Kwiatkowski, DP., and Sabeti, PC. (2014) 'Natural selection and infectious disease'

toxin 'hyperactivates a key signalling molecule in intestinal cells' which increases the severity of the disease.[21] Therefore blood group can affect the severity of cholera.

A further example is found in the human response to HIV. Furthermore, some resistance to HIV has been found for example, a 32-base deletion in CCR5 -–A mutation in the cell surface receptor gene prevents this receptor from being expressed on T cells. T cells are a specific type of lymphocyte involved in both attacking pathogens and in the development of memory cells to help with long term immunity. This mutation has been shown to cause increased resistance to HIV in homozygous individuals.[22]

The human leukocyte antigen (HLA) complex, also known as the major histocompatibility complex (MHC) is 'a group of identification molecules located on the surface of cells that enable the body to distinguish self from non-self'.[23] In the MHC are class I and class II loci that are important for the production of molecules that affect the antigens on immune cells.[24] It has been found that in populations with more variation in pathogen types, the MHC genes also have more variation.[25] It appears that that larger the variety of pathogens present in an area, the higher the selective pressure. This leads to more variation in immune genes and therefore a greater advantage against a wider range of diseases. A study by Prugnolle also found that HLA class I genes with higher diversity compared to average levels were in populations with higher pathogen diversity, which correlates with this.[26]

The MHC molecules are responsible for binding small protein parts from pathogens on to cell surfaces so that they can be recognised by a suitable T-cell (e.g., T killer cell).[27] The MHC has two important characteristics which are that it is both polygenic and polymorphic. This means that within an individual there are lots of different genes for protein binding and that within a population there are lots of different alleles of genes.[28] Cagliani and Sironi stated that the 'high level of diversity at MHC genes is the result of both balancing and directional selection.' Furthermore, level of diversity (polymorphism) in the MHC genes show that it has adapted to pathogens which has resulted in 'changes in the amino acids in the antigen-binding groove.'[29] This means that the MHC is a hugely important part of immunity both individually and on a population scale.[30]

Natural selection is not just exclusive to the MHC and genes responsible for antigen processing. There is a gene that codes for a restriction factor (anti-viral proteins that restrict viral replication[31]) for influenza A virus. This gene protects individuals from influenza and can affect 'protein availability and severity of infection.' It has also been found to limit other viral infections which suggests that that particular gene has evolved as a result of viruses specifically.[32]

---

[21] Bhandari T. (2016) *Study may explain why people with type O blood more likely to die of cholera.* Available at: Study may explain why people with type O blood more likely to die of cholera – Washington University School of Medicine in St. Louis (wustl.edu) (Accessed: 5 March 2022)

[22] Karlsson, EK., Kwiatkowski, DP., and Sabeti, PC. (2014) 'Natural selection and infectious disease'

[23] Delves, P J. (2021) 'Human Leukocyte Antigen (HLA) System', *MSD Manual.* Available at: Overview of the Immune System - Immune Disorders - MSD Manual Consumer Version (msdmanuals.com)

[24] Cagliani, R. and Sirona, M. (2013) 'Pathogen-Driven Selection'

[25] Barreiro, L. and Quintana-Murci, L. (2010) 'From evolutionary genetics'

[26] Cagliani, R. and Sirona, M. (2013) 'Pathogen-Driven Selection'

[27] Janeway, CA., Travers, P., and Walport, M, et al. (2001) *Immunobiology: The Immune System in Health and Disease. 5.* New York: Garland Science. *The major histocompatibility complex and its functions.* Available at: The major histocompatibility complex and its functions - Immunobiology - NCBI Bookshelf (nih.gov)

[28] Janeway, CA., Travers, P., and Walport, M, et al. (2001) *The major histocompatibility complex*

[29] Cagliani, R. and Sirona, M. (2013) 'Pathogen-Driven Selection'

[30] Janeway, CA., Travers, P., and Walport, M, et al. (2001) *The major histocompatibility complex*

[31] No author (2022) Nature Portfolio, *Restriction Factors.* Available at: Restriction factors - Latest research and news | Nature (Accessed: 5 March 2022)

[32] Cagliani, R. and Sirona, M. (2013) 'Pathogen-Driven Selection'

Conclusion

To conclude, pathogens have shaped the human genome largely through natural selection. The impact of pathogens on the genome is influenced by factors such as pathogen exposure and migration. There are however significant difficulties involved when trying to identify different signatures of natural selection. Malaria is an example of a pathogen which affects the human genome, and which leads to mutations in genes involved in red blood cells. Mutations in MHCs also demonstrate the influence of pathogens on the human genome. As a final thought, it would be interesting to compare how the genomes of different species have been affected differently by a pathogen. This could also benefit animal medicine, as well as clinical trials for human medicine. It would be interesting to investigate the different responses to disease, both through history, and when infected.

Bibliography

Barreiro, L. and Quintana-Murci, L. (2010) 'From evolutionary genetics to human immunology: how selection shapes host defence genes', *Nature Reviews Genetics,* 11, pp 17-30. Available at: From evolutionary genetics to human immunology: how selection shapes host defence genes | Nature Reviews Genetics

Bhandari T. (2016) *Study may explain why people with type O blood more likely to die of cholera.* Available at: Study may explain why people with type O blood more likely to die of cholera – Washington University School of Medicine in St. Louis (wustl.edu) (Accessed: 5 March 2022)

Brinkworth, J F. (2017) 'Infectious Disease and the Diversification of the Human Genome', *Human Biology,* 89(1), pp 47-65. Available at: Infectious Disease and the Diversification of the Human Genome on JSTOR

Cagliani, R. and Sirona, M. (2013) 'Pathogen-Driven Selection in the Human Genome', *International Journal of Evolutionary Biology,* 2013. Available at: Pathogen-Driven Selection in the Human Genome (hindawi.com)

Delves, PJ. (2021) 'Human Leukocyte Antigen (HLA) System', *MSD Manual.* Available at: Overview of the Immune System - Immune Disorders - MSD Manual Consumer Version (msdmanuals.com)

Gleichmann N. (2020) 'Innate vs Adaptive Immunity' *Technology Networks, Immunology and Microbiology.* Available at: Innate vs Adaptive Immunity | Technology Networks

Janeway, CA., Travers, P., and Walport, M, et al. (2001) *Immunobiology: The Immune System in Health and Disease. 5.* New York: Garland Science. *The major histocompatibility complex and its functions.* Available at: The major histocompatibility complex and its functions - Immunobiology - NCBI Bookshelf (nih.gov)

Karlsson, EK., Kwiatkowski, DP., and Sabeti, PC. (2014) 'Natural selection and infectious disease in human populations' *Nature Reviews Genetics,* 15, pp 379-393. Available at: Natural selection and infectious disease in human populations | Nature Reviews Genetics

No author (2022) Nature Portfolio, *Restriction Factors.* Available at: Restriction factors - Latest research and news | Nature (Accessed: 5 March 2022)

Metzler, K. (2021) *Selective pressure.* Available at: Selective Pressure: Definition & Example - Video & Lesson Transcript | Study.com (Accessed: 10 March 2022)

Ovsyannikova, IG., Haralambieva IH., Crooke, SN., Poland, GA., Kennedy, RB (2020) 'The role of host genetics in the immune response to SARS-CoV-2 and COVID-19 susceptibility and severity' *Immunological Reviews.* Available at: The role of host genetics in the immune response to SARS-CoV-2 and COVID-19 susceptibility and severity (wiley.com)