Abstract

It was first recognised in 1833 that there was an inherited predisposition to Dupuytren’s Disease, before DNA and genes were discovered (1). Since then, hand surgeons have continued to be fascinated by the nature and extent of this predisposition. As epidemiological methodology and molecular genetic techniques have developed, our understanding has advanced, such that the design of novel therapeutic interventions based on this knowledge is becoming a realistic possibility.

In this section, we begin by briefly reviewing important principles and concepts of genetic inheritance. We then assess the evidence for a genetic predisposition to Dupuytren’s Disease. Next we consider the molecular genetic studies that have helped define the magnitude and nature of that genetic predisposition. Finally, we consider potential future developments in this fast developing field.

Genetics of Dupuytren's disease

The basics of genetic inheritance

Introduction

Genetic information is passed on from generation to generation via the DNA. DNA is a complex molecule organised into the famous double helix configuration, made up of two twisting paired strands. The information passed on from generation to generation is encoded in the bases of this complex molecule - named Guanine, Cytosine, Adenine, and Thymine (G, C, A, T) - and it is the order of these bases along the DNA molecule that makes up the human genome sequence. Each base is paired with a consistent partner on the other strand of DNA - A pairs with T, and G pairs with C. The DNA is organised into chromosomes, and each of us has 22 pairs of “autosomes”, and either two X chromosomes or an X chromosome and a Y chromosome, which determine if we are female or male respectively. We inherit one copy of each chromosome from our mother and one from our father at fertilisation, and our complete genome is present in all of the cells of our body, except spermatozoa and oocytes. This means that we each have two copies (called alleles) of every gene, one from each parent.

In total, the human genome consists of around three billion base pairs. Some of this DNA sequence is transcribed and translated within our cells to make proteins. These regions of DNA are called protein-coding genes. In total, humans have around 24,000 protein coding genes. In fact, these genes only make up around 1.5% of the entire DNA contained in our cells. The rest of the genome consists of “non-coding” DNA. This non-coding DNA has important functions, including the highly complex spatial and temporal regulation of expression of protein coding genes.

Before DNA was discovered, an Austrian monk named Gregor Mendel established the principles of inheritance of genes [2]. He experimented on pea plants, cross-fertilising those with opposing characteristics - tall with short, smooth with wrinkled - and also their offspring. His results established the law of segregation, that there are dominant and recessive traits passed on from parent to offspring (see Box 1). As such, he is considered to be the founding father of modern genetics.

<table>
<thead>
<tr>
<th>Box 1: Definitions used in Mendelian Inheritance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Autosomes</strong></td>
</tr>
<tr>
<td>The non-sex chromosomes, that is chromosomes 1–22.</td>
</tr>
<tr>
<td><strong>Sex chromosomes</strong></td>
</tr>
<tr>
<td>The X and Y chromosomes. The Y chromosome is much smaller than the X chromosome, so the vast majority of genes</td>
</tr>
</tbody>
</table>
present on the X chromosome are not present on the Y chromosome. A normal female carries two X chromosomes, and therefore has two copies of each gene on the X chromosome. A normal male has one X and one Y chromosome, so has only one copy of most genes present on the X chromosome.

**Autosomal dominant**

When a single mutant copy of a gene is present, and the effects of that mutation are seen in the individual. The gene in question is present on one of the 22 autosomes. These diseases are passed from generation to generation with a 50% chance that the offspring of an affected individual will also be affected.

**Autosomal recessive**

Mutations in both copies of the gene are required for the effects of the mutation to be seen in the individual. The gene in question is present on one of the 22 autosomes. These diseases usually occur "sporadically" within a family, rather than being passed on from generation to generation. It is more likely to be found in the offspring of consanguineous mating. Unaffected individuals with one mutant copy of a gene are known as carriers.

**X-linked recessive**

In males, mutations in their single copy of a gene present on their only X chromosome is sufficient to cause the disease. In females, mutations in both copies of the gene present on their two X chromosomes would be required to produce the disease. For this reason, X-linked recessive traits are much more common in males. Affected fathers cannot pass on these diseases to their sons, as they pass on their Y chromosome to their sons. All daughters of affected fathers will be unaffected carriers of the disease.

**X-linked dominant**

In both males and females, a single mutant copy of a gene on the X chromosome is sufficient to cause the disease. Again, affected fathers cannot pass on these diseases to their sons, as they pass on their Y chromosome to their sons. However, all daughters of affected fathers will have the disease.

What is normal variation, and what is a mutation?

Within the three billion base pairs of our genome, around 30 million base pairs are what is known as polymorphic. This means they exist in two or more forms within the normal population. This is the normal variation that makes each of us different from the other. A minority of these variants are within protein coding genes, and will have a direct effect of the function of that gene. However, the vast majority of these variants are within non-coding DNA, and have subtle effects within the cell, such as changing the expression of a protein-coding gene. Some of these effects will be beneficial, and some deleterious, to the survival and reproductive fitness of the individual. Such variants will be retained or lost, respectively, within the population over time. This is the basis of Darwin’s theory of natural selection.

In genetic terms, a mutation is a change in the DNA sequence. This may take the form of a change in a single base pair, or the deletion, insertion, or rearrangement of larger parts of DNA, or even a whole chromosome. A mutation generally has the property of being disadvantageous to the organism. Most mutations are found within the protein coding genes, and lead to disruption of either the whole of the protein that the gene encodes, or a vital part of the protein required for its function. Diseases caused by such mutations tend to be inherited in regular predictable patterns - Autosomal Dominant, Autosomal Recessive, and X-linked recessive, see Box 1 - and are termed Mendelian diseases in honour of Gregor Mendel. Because the effects of these mutations are deleterious, their frequency within the population is usually very low. Furthermore, the syndromes caused by such mutations are usually (though not exclusively) present and recognised at birth or in childhood.

Mendelian diseases within hand surgery

Hand surgery occupies a special place in the history of human genetics, as the very first human disease shown to follow an autosomal dominant Mendelian pattern of inheritance was brachydactyly type A1. Subsequently, this has been shown to be caused by mutations in the gene Indian Hedgehog (IHH) [3]. Most congenital hand malformation syndromes display autosomal dominant inheritance. A rare example of an autosomal recessive hand surgery condition is one of the variants of split-hand/split foot malformation, characterised by central clefts of the hands and feet. This has been shown to be caused by mutations in the gene WNT10B [4-7]. X-linked recessive congenital hand malformations are extremely rare, an example being...
a family with fusion of the 4th and 5th metacarpals [8].

However, Mendelian inheritance is not always clear-cut. For example, not all people who carry a particular dominant mutation will show signs of the disease on clinical examination. This phenomenon is known as incomplete penetrance. For example, a mutation within a regulatory region of the gene Sonic Hedgehog has been shown to cause triphalangeal thumb [9]. In this study, only around half of people who carried the mutation were found to actually have triphalangeal thumb on clinical examination. Also, patients with precisely the same mutation in the same gene may have a slightly different phenotype. This is known as variable expressivity. These phenomena are due to the modifying effects of other variants within the genome, the effects of the environment, or to random effects.

**What is a complex disease?**

As a consequence of the above, it is helpful to consider variants and mutations as lying on a spectrum (Figure 1). At the one end are mutations that are very rare, but have a very large increased risk of causing a particular disease within an organism. At the other end are variants that are very common, but have a very small increased risk of causing a particular disease within an organism. For Mendelian diseases, only rare mutations in a single gene with a high effect size cause the disease.

However, there are many diseases, for example heart disease and rheumatoid arthritis, whereby a familial predisposition is well recognised, but mutations in single genes are not responsible for this predisposition, and the inheritance does not follow simple Mendelian patterns. Here, multiple common variants each slightly increase the risk of a person getting the disease. These variants are often in non-coding areas of the DNA, and exert their effects by subtle changes in gene function. Furthermore, they interact with environmental influences, and this interaction determines the final expression of the disease. These diseases are known as complex diseases, and discovering the basis of the genetic predisposition is more difficult than for Mendelian diseases.

![Figure 1: The spectrum of genetic mutations and variations. Traditional mutations are rare, and have a large risk of affecting the phenotype (top left). Common variants have a much smaller risk of affecting the phenotype, but multiple alleles of small effect can combine to produce a complex disease. Rare variants are between the two extremes, and might pre-dispose to complex disease in a small proportion of the population.](image-url)

When considering complex diseases, it is helpful to think of a threshold effect (Figure 2). Here, many genetic variants each increase a person’s chance of getting a particular disease by a small amount. The more of these predisposing variants that an
individual carries, the closer they are to tipping over the threshold, and displaying the disease. Environmental factors interact with these genetic factors to provide a further push, leading to the presence of the disease in an individual. If there is a greater genetic predisposition (a stronger family history), more predisposing genetic factors are present, and less environmental factors are required to cause the disease to become manifest. Similarly, a person with little genetic predisposition can still be pushed over the threshold and express the disease by having multiple excessive environmental exposures.

![Figure 2: The threshold model of genetic predisposition to complex disease. Blue boxes represent genetic predisposition, and red boxes represent environmental predisposition. The horizontal black line represents a threshold, below which the individual is clinically unaffected. Patient A has a low genetic predisposition and low environmental exposure, and does not have the disease. Patient B has a large genetic predisposition, and a low environmental exposure is enough for them to exceed the threshold and be clinically affected. Patient C has a low genetic predisposition, but a large environmental exposure, and is also clinically affected. Finally, Patient D has a large genetic and environmental exposure, and is more severely affected.]

A further consequence of this model, is that people with a larger genetic predisposition tend to have earlier onset disease - they need less environmental insult to push them over the threshold - and the disease tends to be more severe - they get further over the threshold with a similar amount of environmental exposure to a person who has a smaller genetic predisposition.

When considering this threshold model, it becomes clear that Dupuytren’s disease is very typical of a complex disease, just like heart disease, diabetes, and rheumatoid arthritis. In Dupuytren’s disease we observe a condition that has a familial tendency, but doesn’t follow simple Mendelian inheritance rules. The onset of the disease is in later life, and earlier onset disease tends to have a stronger family history than later onset disease. Patients with a strong family history also tend to be more severely affected. Furthermore, there are clear and well documented environmental influences on the expression of the disease. As shown towards the end of the chapter, this concept of Dupuytren’s disease being a complex disease has proved to be correct, and we are now beginning to understand some of the genetic variants that cause the predisposition. We hope that this will form the basis for the rational design of new therapies aimed at preventing both the disease in the first place, and also recurrence after treatment in high-risk individuals.

The evidence for a genetic predisposition to Dupuytren’s disease

Population studies

Modern genetic studies have defined the frequency of variants throughout the genome in different populations. This has confirmed the hypothesis that populations from different parts of the world have distinguishable genetic backgrounds. If a disease is more common in particular racial groups, then a genetic predisposition to that particular disease is possible. Many studies have evaluated the prevalence of Dupuytren’s disease within a particular country (reviewed in [10]). Estimates of prevalence vary widely depending on the assessor, diagnostic criteria, and co-morbid conditions of the population under study.

When considering studies not designed to evaluate a cohort with a specific risk factor for Dupuytren’s disease, the prevalence in Scandinavian countries varies from 6% to 11% [11-13]; in Iceland it is reported to be between 13% and 19% [14, 15]; in the
UK 3.5% to 30% [16-21]; in France and Spain 8% to 9% [22, 23]; in North America, the prevalence is 3.4% to 33% [24, 25].

In the Netherlands, the overall community prevalence was 22.1%, but those with contracture were only 4.2% of the study population [26]. Similarly, in Belgium, the population prevalence was 32%, but only 8% of males and 4% of females had contracture [27]. These results suggest a generally benign natural history of the disease, though population-based longitudinal studies would be required to confirm this. Interestingly, the documented prevalence in India is also 4% [28]. Dupuytren’s disease among black Africans is limited to case reports [29].

Of course, populations also share environmental influences, so stronger evidence that genetics is important comes from studies that demonstrate a different incidence in different racial groups within the same environment, or demonstrate that the incidence of a disease persists when the population moves to a different environment. Finsen and co-authors found that the prevalence of Dupuytren’s disease amongst the native Sami of northern Scandinavia was 5.8%, and in Norwegians living in the same area was 10.1% [30]. Though this difference was not statistically significant, it does suggest that genetically distinct populations have different prevalence of disease. Furthermore, Heuston studied the prevalence of Dupuytren’s disease in Australia, a country populated by people from northern Europe who have migrated to a different environment [31]. He found a prevalence rate of 23%, equivalent or higher than that found in Europe.

Overall, Dupuytren’s is widely considered to be a disease of Europeans, and more specifically Northern Europeans, but large variability in the data make this conclusion less secure than is generally assumed. Detailed genetic studies should provide a more definitive answer to this question in the coming years.

**Twin and family studies**

Studying twins is a powerful way to document and quantify the magnitude of genetic risk in a disease. In twin studies, concordance for a particular disease is defined as twins having the same disease status - i.e. both twins have the disease or both twins do not have the disease - whereas discordance is defined as twins not having the same disease status - i.e. one twin has the disease and the other does not. Monozygotic (identical) twins share their entire genome, whereas dizygotic (non-identical) twins, like normal siblings, share on average 50% of their genome. Thus, if monozygotic twins are more commonly concordant for a particular disease than dizygotic twins, this is good evidence that genetics play a part in the predisposition to disease.

In Dupuytren’s disease, no systematic review of twin registries has been published to date. The data on twin concordance is limited to case reports, and shows variable effects, with some authors reporting concordant twin pairs [32], and others discordant twin pairs [33, 34].

Several authors have attempted to document and quantify the familial incidence of Dupuytren’s disease. In the UK, 7-15% of those with Dupuytren’s disease reported a family history of the disease [35, 36], and in France the corresponding figure was 20% [23]. However, many people with early Dupuytren’s disease may not recognise that they are affected, and certainly their relatives may be unaware of them having the condition. It is therefore likely that these studies underestimate the true incidence of a positive family history. Indeed, Ling obtained a positive family history of Dupuytren’s disease in 34/50 patients but discovered that 34/50 patients actually had a relative with Dupuytren’s disease when he personally examined 830 of their relatives [37].

A standard method of quantifying familial risk to disease is the sibling recurrence risk. This compares the rate of the disease under study in the siblings of index cases, compared to the rate in the general population. Because siblings share on average 50% of their DNA, a higher prevalence in siblings is strong evidence for a genetic predisposition. For example, in rheumatoid arthritis, a sibling recurrence risk of between 3 and 7 has been reported, with higher values for more severe disease (38), and in anteromedial osteoarthritis of the knee, the sibling recurrence risk has been calculated to be 3.21 [39].

In Dupuytren’s disease, the sibling recurrence risk has been calculated in the UK. The authors examined a randomly selected single sibling of 100 Dupuytren’s disease index cases, and compared the rate of Dupuytren’s disease to that of 124 controls ascertained from an ophthalmology clinic for non-diabetic eye conditions. They found that 47% of siblings were affected compared to 10.5% of controls, giving a sibling recurrence risk of 4.5 [40]. Furthermore, using a history of surgery for Dupuytren’s disease as a proxy marker of disease severity, they found the sibling recurrence risk for surgery to be 7.1, suggesting more severe disease is more highly heritable, consistent with the threshold model of complex disease presented above (figure 2).

**Molecular studies**
Linkage studies

In single gene disorders, large families where a disease is passing from generation to generation can be used to determine where on the chromosomes a genetic mutation using molecular techniques of genetic linkage. Here, genetic markers are typed across the entire genome in affected and unaffected family members, and segments of chromosomes where the variants are similar in affected individuals, but different in unaffected individuals are said to be “linked” to the disease. This technique works well for Mendelian disease, where rare mutations have a large effect (figure 2) and for decades was the bedrock for discovering disease-causing mutations for several decades. However, for complex diseases linkage cannot detect common variants that predispose to disease [41, 42].

It is possible, in a small minority of cases, that a rare variant in a gene is predisposing to disease with a high effect size in a single family, or several closely related families, effectively acting like a mutation. Such predisposition could be studied by linkage analysis, and there is a single report of this in the literature for Dupuytren’s disease [43]. Here, researchers performed a linkage analysis on a large five-generation Swedish family, and found evidence of linkage to a region on chromosome 16 that just reached statistical significance. This result has not been confirmed since the original publication.

Association studies

Genetic association studies test for a correlation between disease status and genetic variants in cases and controls. A higher frequency of an allele (see introduction) in cases compared to controls suggests that the tested variant, or another untested variant that is close to the tested variant, predisposes to disease.

There are two important principles in association studies. Firstly, any variant found to be associated with a complex disease should be replicated in an independent cohort of cases and controls. Secondly, when multiple different variants are tested for association, the p-value accepted as statistically significant should be appropriately lowered by a Bonferroni correction [44]. These measures help to guard against statistical false positives.

Association studies can be divided into candidate gene studies and genome-wide association studies (GWAS).

Candidate gene studies

Before the advent of molecular genetic techniques that enabled researchers to study variation across the entire genome, researchers hypothesized which of the approximately 24,000 protein-coding genes might be involved in a particular complex disease. They found a small minority of the hundreds of variants within the gene, and tested them for association with the disease under study, typically using tens or a few hundred cases and controls. As will be evident from the above discussion of human variation, the chance of correctly identifying a predisposing variant in this way is small.

Furthermore, experience from GWAS (see below), has demonstrated that the majority of potential candidate gene associations reported across the literature in a wide variety of complex diseases, do not stand up to rigorous independent replication [45-47].

A similar picture has emerged with Dupuytren’s disease. Associations between variants in many different genes have been investigated in small sample groups. These studies often also have methodological flaws, such as reporting statistical significance without performing a Bonferroni correction for multiple hypothesis testing [48]. Some positive associations have been reported [48-53]. Some negative associations have also been reported [54-57]. Of note, none of the positive associations have been replicated in independent cohorts, or in the GWAS described below.

Genome-wide association studies (GWAS)

The current gold-standard technique for studying the genetics of complex disease is the GWAS. Here, hundreds of thousands or millions of variants across the entire genome are typed in cases and controls, and tested for association with the disease. The biggest advantage of this technique is that it does not require any pre-existing knowledge of the pathogenesis of the disease in question: all areas of the genome are interrogated for association with the disease and assessed equally. This has led to the discovery of associations of genes with diseases that were previously unsuspected, leading to new avenues of research into disease pathogenesis. Indeed, the majority of robust associations that have been discovered are not within protein-coding genes at all, but non-coding DNA, and are likely to regulate gene expression [58].

However, the vast number of statistical tests performed in a GWAS leads to an unprecedented potential for false-positive results. Thus an adequate Bonferroni correction must be performed, typically leading to $P < 0.00000005$ being accepted as
statistical evidence for association. This means that thousands of cases and controls must be tested for association in order to achieve significance [59, 60].

The GWAS approach has proved extremely successful in the search for variants predisposing to complex disease, with over 15,000 associations reported in the National Human Genome Research Institute catalogue of published GWAS [61]. Whilst the direct application of these results to clinical medicine has proved to be frustratingly slow [62], progress towards clinical application is being made in diseases such as rheumatoid arthritis [63].

In Dupuytren’s disease, a European GWAS has been reported [64]. In phase 1 of the study, the authors genotyped 960 Dutch patients with Dupuytren’s disease, and 3,117 controls patients in a discovery cohort. They then attempted replication of the significant results from phase 1 in an independent replication cohort consisting of 1,365 cases and 8,445 controls from the UK, Germany and Holland. The results of this analysis revealed nine genetic regions that pre-dispose to Dupuytren’s disease. None of these associated variants were in protein-coding genes, and none of them were close to any of the reported associations from previous candidate gene association studies.

The authors went on to perform a bioinformatic analysis of the protein coding genes that were close to the variants associated with Dupuytren’s disease. This revealed a significant excess of genes involved in the WNT signaling pathway. This suggests that WNT signaling may be important in the pathogenesis of Dupuytren’s disease, though this observation requires further investigation.

The results of this GWAS have been used to further explore the genetics of Dupuytren’s disease. Calculations based on the sibling recurrence risk reveal that the genetic regions discovered in this study account for around 12% of the genetic predisposition to Dupuytren’s disease, implying that larger studies with more cases and controls will discover many more associated variants [40]. Furthermore, a positive correlation has been described between increasing numbers of genetic risk variants and early onset disease, positive family history, and features of a Dupuytren’s diathesis [65]. Again, this data is consistent with the threshold model presented in figure 2. A further independent study has confirmed that the most statistically significantly associated variant is associated with Dupuytren’s disease in a Polish population [66].

Future perspectives

Further genetic studies

Molecular genetic technology continues to evolve at a rapid pace. It is evident from studies of other complex diseases that studying larger case-control cohorts leads to the discovery of further common variants that pre-dispose to disease with smaller effect size [67], and these variants can give valuable insights into the biology of the disease. This principle is currently being applied to Dupuytren’s disease, with a much larger GWAS currently being undertaken in European populations.

When the first draft of the human genome was published in 2001, it was estimated that the cost of this sequence was around $100 million. It is now possible to sequence an entire human genome for under $7,000, and the cost is still falling [68]. We expect that genetic sequence data will translate into routine clinical practice in many areas over the coming decades. We therefore also expect that Dupuytren’s disease patients will undergo whole genome sequencing, and their sequence data will be compared to controls without Dupuytren’s disease, in order to discover the whole range of genetic variation, both common and rare variants, that predispose to disease. As an intermediate step, it is possible to sequence just the protein-coding genes within the genome – so-called exome sequencing. This technique has proved excellent for discovering mutations causing Mendelian disease, and may also find some application in complex diseases [69, 70].

However, there are currently massive barriers to the storage, analysis, and interpretation of both exome sequencing and whole-genome sequencing data. The development of bioinformatic techniques has lagged behind molecular developments, and training researchers in this area will be the key to further advances in the field.

Implications for treatment

The goal of genetic research into Dupuytren’s disease is to understand the pathophysiology of the disease, leading to the rational design of new treatments aimed at preventing development of the disease, and preventing recurrence after surgical or other treatment. Furthermore, in the future it is possible that genetic risk data will be taken into account when planning surgery. For example, a patient with a high risk of recurrence based on genetic and clinical data might require a more extensive primary procedure.
These goals are realistic in the medium to long term, but many challenges remain. As surgeons, we have unique access to patients with this fascinating disease. We should strive to improve the treatment of current patients by recruiting to high quality randomized controlled trials of current treatments, and strive to improve the treatment of future patients by recruiting to high quality basic and applied science research into the disease.

References

61. Hindorff LA, JEBI, Morales J (European Bioinformatics Institute), Junkins HA, Hall PN, Klemm AK, and Manolio TA. A Catalog of Published Genome-Wide Association Studies.
68. KA W. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP) Available at: www.genome.gov/sequencingcosts. Accessed [03/02/2014].

Citation Note

Doi of this Chapter
10.5680/lhhs000010

This Genetics of Dupuytren's disease, Genetics of Dupuytren's disease, is licensed under the Creative Commons Attribution-ShareAlike 4.0 International license, although certain works referenced herein may be separately licensed.