# Exploring Genetics Through Genetic Disorders

*Teacher Guide*

## *Assigning Allele Profiles*

Within each disorder, the alleles are numbered so that the **lower numbers are easiest**, and the difficulty goes up with higher numbers. The alleles for **hemophilia and hemoglobin disorders are more challenging** than the rest.

All the alleles are based on real ones described in the literature (references are listed at the end). We changed the names of the alleles mainly so we could streamline the text for students. The tables below include the names that are used by the scientific community.

### Alpha-1 Antitrypsin Deficiency

| Allele | Type of change | Difficulty & notes |
|---|---|---|
| D1 – M (Mineral Springs) | Single-base substitution | Easy |
| D2 – W (Bethesda) | Single-base substitution | Easy |
| D3 – Z | Single-base substitution | Medium |
| D4 – M (Malton) | Three-base deletion | Medium |
| D5 – NULL (Mattawa) | Single-base insertion | Medium; the DNA/amino acid data is a little harder to interpret |
| D6 – Pittsburgh | Single-base substitution | Bonus/Advanced |

The Z allele (aka D3) is by far the most common disorder-causing allele of SERPINA1. Another common allele, S, is not included in this activity. The S allele codes for a version of AAT protein with reduced function that differs from the healthy protein by one amino acid. The S allele is fairly common, but it generally causes AAT deficiency only when it is in combination with a Z or null allele.

### Cystic Fibrosis

The Demo Lab Notebook is filled in with information for allele C1. If you use the Demo, students may have an easier time filling in the information for the other cystic fibrosis alleles as well.

| Allele | Molecular changes | Difficulty |
|---|---|---|
| C1 – G542X | Single-base substitution | Easy; don't assign if using the Demo Lab Notebook |
| C2 – F508del | Three-base deletion | Easy |
| C3 – G551D | Single-base substitution | Easy |
| C4 – R1070W | Single-base substitution | Medium |
| C5 – A455E | Single-base substitution | Medium |

The F508del allele is the most common disorder-causing allele of CFTR. It's present in >70% of people with the disorder. Most of the work done on R1070W and A455E is done with F508del as the second allele. To make the data more understandable, we estimated the sweat chloride levels for R1070W / R1070W and A455E / A455E, based on papers that studied protein function.

## Hemoglobin Disorders

These disorders are a little on the tricky side. Variations in the HBB gene can cause several distinct disorders, and the symptoms and molecular mechanisms vary widely. To understand their alleles, students will need to read and process more information.

| Allele | Molecular changes | Difficulty |
|---|---|---|
| HB1 – HbS | Single-base substitution | Medium |
| HB2 – HbC | Single-base substitution. | Medium |
| HB3 – Glu6FS | Single-base deletion | Medium |
| HB4 – HbE | Single-base substitution | Medium+ |
| HB5 – E121 to TER | Single-base substitution | Advanced |
| HB6 – Hemoglobin Denver | Single-base substitution | Advanced |

HbS is the most common allele in sickle cell disease. HbC and HbE, because they also contribute to malaria resistance, are also quite common. The other alleles are rare.

## Hemophilia

This disorder is a little tricky. It takes a little work to understand how the proteins interact with others to help blood clot, and how the variations affect those interactions.

In the literature, *F8* and *F9* alleles are referred to by a code that indicates the position of the change in the amino acid sequence and the type of change. The amino acid numbering system changed around the year 2000. We use the current system, though some publications still use "Legacy" numbering. See reference section for details.

| Gene, Allele | Molecular changes | Difficulty and notes |
|---|---|---|
| H1 – *F8*, A415V | Single-base substitution | Easy |
| H2 – *F9*, R449Q | Single-base substitution | Easy |
| H3 – *F9*, L19F*1 | Single-base deletion | Easy |
| H4 – *F8*, A303E | Single-base substitution | Medium |
| H5 – *F9*, F55I | Single-base substitution | Medium. |
| H6 – *F8*, intron 22 inversion | Chromosomal inversion. A piece of chromosome broke away, rotated 180 degrees, and fused back in place. | Bonus/Advanced. The molecular changes are quite different from those for the other alleles. There's a lot to read & process about the rearrangement, and it may be helpful to have some knowledge of gene regulation. |

The H6 allele is the most common allele in hemophilia. It's found in nearly 50% of all severe cases. Many of the other hemophilia alleles are the result of "founder effect," and a specific allele can be traced back to a common ancestor. In different parts of the world, new mutations occurred, creating new alleles that were passed on. This is why different hemophilia alleles tend to be more common in people with ancestors from different places. For example, the H2 allele is one of the most common alleles in people with French ancestry, but it is rare in other populations.

### Marfan Syndrome

Autosomal dominant inheritance pattern. In the literature, *FBN1* alleles are usually referred to by a code that indicates the position of the change in the amino acid sequence and the type of change.

| Allele | Molecular changes | Difficulty |
|---|---|---|
| M1 – G1013R | Single-base substitution | Easy |
| M2 – C2686F | Single-base substitution | Easy |
| M3 – I1892X | Five-base insertion (ACACT) | Medium |
| M4 – R2726W | Single-base substitution | Medium |
| M5 – C1564S | Single-base substitution | Medium |

Most people with Marfan syndrome have an allele that is unique to their family. In fact, only about 10% of alleles are shared by another family. With the exception of neonatal Marfan syndrome, few connections have been made between the type of allele a person has and the severity of the symptoms they experience.

## *Demo Lab Notebook*

A Demo Lab Notebook is provided as a pdf with information filled in for Cystic Fibrosis allele C1. We suggest projecting the demo lab notebook and showing students how to fill in each section.

## *Notes on mRNA and protein sequences*

Students will get the most from this unit if they have the correct information in the Mutation & Alleles section of their Lab Notebooks. Please print the correct sequences from the answer key and give each student a strip of paper with the information for their allele. They can use it to check their answers and correct any errors.

Here are places where students commonly run into trouble:

- Transcribing the wrong DNA strand: Make sure students copy the template strand (printed in darker text, with upside-down letters)
- Transcribing in the wrong direction: Make sure students go from left to right as they both read the DNA template and write the mRNA sequence.
- Misreading the Amino Acid Coding chart: You may want to go over this with students before they begin this section.

# *References*


### *Alpha-1 Antitrypsin Deficiency*

Brode, S. K., Ling, S. C., & Chapman, K. R. (2012). Alpha-1 antitrypsin deficiency: a commonly overlooked cause of lung disease. *Canadian Medical Association Journal, 184*(12), 1365-1371

Crystal, R.G. (1990). Alpha 1-antitrypsin deficiency, emphysema, and liver disease. Genetic basis and strategies for therapy. *The Journal of clinical investigation*, 85(5), 1343-1352.

Ferrarotti, I., Thun, G. A., Zorzetto, M., Ottaviani, S., Imboden, M., Schindler, C., ... & Probst-Hensch, N. M. (2012). Serum levels and genotype distribution of α1-antitrypsin in the general population. *Thorax, 67*(8), 669-674.

Hua, B., Fan, L., Liang, Y., Zhao, Y., & Tuddenham, E. G. (2009). α1-antitrypsin Pittsburgh in a family with bleeding tendency. haematologica, 94(6), 881-884.

Lewis, J. H., Iammarino, R. M., Spero, J. A., & Hasiba, U. (1978). Antithrombin Pittsburgh: an alpha1-antitrypsin variant causing hemorrhagic disease. *Blood, 51*(1), 129-137.

Owen, M. C., Brennan, S. O., Lewis, J. H., & Carrell, R. W. (1983). Mutation of antitrypsin to antithrombin: α1-antitrypsin Pittsburgh (358 Met –> Arg), a fatal bleeding disorder. *New England Journal of Medicine, 309*(12), 694-698.

Information about the SERPINA1 alleles came from the following sources (accessed March 2018):

- Online Mendelian Inheritance in Man (OMIM), entry 107400
- DNA and amino acid sequences were accessed through UniProt, entry P01009

  *Note that the first 24 amino acids make up a signal peptide that is later cleaved to make the mature protein. Some sources number the amino acids according to their position in the mature protein (excluding the signal peptide). We have numbered them here according to their position relative to the translation start codon (including the signal peptide).*

### *Cystic Fibrosis*

Bompadre, S. G., Sohma, Y., Li, M., & Hwang, T. C. (2007). G551D and G1349D, two CF-associated mutations in the signature sequences of CFTR, exhibit distinct gating defects. *The Journal of General Physiology, 129*(4), 285-298.

Cebotaru, L., Rapino, D., Cebotaru, V., & Guggino, W. B. (2014). Correcting the cystic fibrosis disease mutant, A455E CFTR. *PloS One, 9*(1), e85183.

Farrell, P. M., & Koscik, R. E. (1996). Sweat chloride concentrations in infants homozygous or heterozygous for F508 cystic fibrosis. *Pediatrics*, 97(4), 524-528.

Krasnov, K. V., Tzetis, M., Cheng, J., Guggino, W. B., & Cutting, G. R. (2008). Functional studies of rare missense mutations in CFTR facilitate interpretation of genotype-phenotype relationships. *Human Mutation, 29*(11), 1364.

Will, K., Dörk, T., Stuhrmann, M., Hardt, H. V. D., Ellemunter, H., Tümmler, B., & Schmidtke, J. (1995). Transcript analysis of CFTR nonsense mutations in lymphocytes and nasal epithelial cells from cystic fibrosis patients. *Human Mutation, 5*(3), 210-220.

Information about the *CFTR* alleles came from the following sources (accessed August 2018):

- https://cftr2.org/
- DNA and amino acid sequences were accessed through UniProt, entry P13569.

***Hemoglobin Disorders***

Forget, B. G., & Bunn, H. F. (2013). Classification of the disorders of hemoglobin. *Cold Spring Harbor perspectives in medicine, 3*(2), a011684.

Gerald, P. S., & Efron, M. L. (1961). Chemical studies of several varieties of Hb M. *Proceedings of the National Academy of Sciences, 47*(11), 1758-1767.

Kohne, E. (2011). Hemoglobinopathies: clinical manifestations, diagnosis, and treatment. *Deutsches Ärzteblatt International, 108*(31-32), 532.

Lohani, N., Bhargava, N., Munshi, A., & Ramalingam, S. (2018). Pharmacological and molecular approaches for the treatment of β-hemoglobin disorders. *Journal of Cellular Physiology, 233*(6), 4563-4577.

Origa, R. (2017). β-Thalassemia. *Genetics in Medicine, 19*(6), 609.

Patrinos, G. P., Giardine, B., Riemer, C., Miller, W., Chui, D. H., Anagnou, N. P., ... & Hardison, R. C. (2004). Improvements in the HbVar database of human hemoglobin variants and thalassemia mutations for population and sequence variation studies. *Nucleic acids research, 32*(suppl_1), D537-D541.

Thein, S. L., Hesketh, C., Taylor, P., Temperley, I. J., Hutchinson, R. M., Old, J. M., ... & Weatherall, D. J. (1990). Molecular basis for dominantly inherited inclusion body beta-thalassemia. *Proceedings of the National Academy of Sciences, 87*(10), 3924-3928.

Information about HBB alleles came from the following sources (accessed January 2019):

- Online Mendelian Inheritance in Man, entry 141900. https://www.omim.org/entry/141900
- HbVar: A Database of Human Hemoglobin Variants and Thalassemias. http://globin.bx.psu.edu/hbvar/menu.html
- DNA and amino acid sequences were accessed through UniProt, entry P68861. https://www.uniprot.org/uniprot/P68871

  *Note that there are two schemes for numbering the amino acids. One counts the first methionine, which is not part of the mature protein, as position one. The other counts the first amino acid in the mature protein as one. We used the former.*

***Hemophilia***

Grant, M. A., Baikeev, R. F., Gilbert, G. E., & Rigby, A. C. (2004). Lysine 5 and phenylalanine 9 of the

factor IX ω-loop interact with phosphatidylserine in a membrane-mimetic environment. *Biochemistry*, 43(49), 15367-15378.

Hakeos, W. H., Miao, H., Sirachainan, N., Kemball-Cook, G., Saenko, E. L., Kaufman, R. J., & Pipe, S. W. (2002). Hemophilia A mutations within the factor VIII A2-A3 subunit interface destabilize factor VIIIa and cause one-stage/two-stage activity discrepancy. T*hrombosis and Haemostasis*, 87(05), 781-787.

Ivaskevicius, V., Jurgutis, R., Rost, S., Müller, A., Schmitt, C., Wulff, K., ... & Oldenburg, J. (2001). Lithuanian haemophilia A and B registry comprising phenotypic and genotypic data. *British Journal of Haematology*, 112(4), 1062-1070.

Kurachi, S., Pantazatos, D. P., & Kurachi, K. (1997). The carboxyl-terminal region of factor IX is essential for its secretion. *Biochemistry*, 36(14), 4337-4344.

Lassalle, F., Marmontel, O., Zawadzki, C., Fretigny, M., Bouvagnet, P., & Vinciguerra, C. (2018). Recurrent *F8* and *F9* gene variants result from a founder effect in two large French haemophilia cohorts. *Haemophilia*, e213-e221.

Rudzki, Z., Duncan, E. M., Casey, G. J., Neumann, M., Favaloro, E. J., & Lloyd, J. V. (1996). Mutations in a subgroup of patients with mild haemophilia A and a familial discrepancy between the one-stage and two-stage factor VIII: C methods. *British Journal of Haematology*, 94(2), 400-406.

Sauna, Z. E., Lozier, J. N., Kasper, C. K., Yanover, C., Nichols, T., & Howard, T. E. (2015). The intron-22–inverted *F8* locus permits factor VIII synthesis: explanation for low inhibitor risk and a role for pharmacogenomics. *Blood*, 125(2), 223-228.

Information about the *F8* and *F9* alleles came from the following sources (accessed December 2018):

- http://www.factorviii-db.org/index.php
- http://www.factorix.org/
- DNA and amino acid sequences were accessed through UniProt, entries P00451 (*F8*) and P00740 (*F9*).

  *Note that the first amino acids make up a signal peptide that is later cleaved to make the mature protein. This stretch is 19 amino acids long for coagulation factor VIII and 46 amino acids long for coagulation factor IX. Some sources use a "Legacy" designation that numbers amino acids starting after the signal peptide. We numbered them according to their position relative to the translation start codon, which includes the signal peptide.*

***Marfan Syndrome***

Jensen, S. A., Aspinall, G., & Handford, P. A. (2014). C-terminal propeptide is required for fibrillin-1 secretion and blocks premature assembly through linkage to domains cbEGF41-43. *Proceedings of the National Academy of Sciences*, *111*(28), 10155-10160.

Jensen, S. A., Iqbal, S., Bulsiewicz, A., & Handford, P. A. (2015). A microfibril assembly assay identifies different mechanisms of dominance underlying Marfan syndrome, stiff skin syndrome and acromelic dysplasias. *Human Molecular Genetics*, *24*(15), 4454-4463.

Kirschner, R., Hubmacher, D., Iyengar, G., Kaur, J., Fagotto-Kaufmann, C., Bromme, D., ... & Reinhardt, D. P. (2011). Classical and neonatal Marfan syndrome mutations in fibrillin-1 cause differential protease susceptibilities and protein function. *Journal of Biological Chemistry*, jbc-M111.

Milewicz, D. M., Grossfield, J., Cao, S. N., Kielty, C., Covitz, W., & Jewett, T. (1995). A mutation in FBN1 disrupts profibrillin processing and results in isolated skeletal features of the Marfan syndrome. *Journal of Clinical Investigation*, *95*(5), 2373-2378.

Raghunath, M., Superti-Furga, A., Godfrey, M., & Steinmann, B. (1993). Decreased extracellular deposition of fibrillin and decorin in neonatal Marfan syndrome fibroblasts. *Human Genetics, 90*(5), 511-515.

Schrijver, I., Liu, W., Odom, R., Brenn, T., Oefner, P., Furthmayr, H., & Francke, U. (2002). Premature termination mutations in FBN1: distinct effects on differential allelic expression and on protein and clinical phenotypes. *American Journal of Human Genetics, 71*(2), 223-237.

Schrijver, I., Liu, W., Brenn, T., Furthmayr, H., & Francke, U. (1999). Cysteine substitutions in epidermal growth factor–like domains of fibrillin-1: distinct effects on biochemical and clinical phenotypes. *American Journal of Human Genetics, 65*(4), 1007-1020.

Information about the FBN1 alleles came from the following sources (accessed August 2018):

- Protein function data for M1 is primarily based off Kirshner 2011. We used data from a number of papers that looked at neonatal Marfan syndrome (such as Raghunath 1993) to estimate the amount of fibrillin-1 protein incorporated into microfibrils.
- Protein function data for M4 combines the patient data from Milewicz 1995 (allele shown) with functional studies performed on a similar allele in Jensen 2014. Both alleles disrupt the same cut site on the fibrillin-1 protein.
- DNA and amino acid sequences were accessed through UniProt, entry P35555

This work was supported by a Science Education Partnership Award (1R25GM021903) from the National Institute of General Medical Sciences of the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.