Case: Gene Regulation of Autism

*(Note: I have greatly simplified and generalized a lot of details to make this case a manageable length and difficulty. If you want to know more, I can point you to good review articles.)*

Overall Background (LMS Pre-class Reading):

We have seen in class how cells can regulate gene activity at multiple points. They include:

1. **Regulating chromatin structure and accessibility.**   
   The structure of chromatin is determined by histone acetylation and methylation, and by DNA methylation.
   1. These modifications determine whether DNA is organized as euchromatin or heterochromatin.
   2. While DNA is in euchromatin form, it is available for use by the cell, but when the DNA in heterochromatin form, it is not available.
   3. DNA is not always one or the other; it can differ:
      1. Between cell types (for example, genes that are in euchromatin in liver cells may be in heterochromatin in neurons), and
      2. At different stages in the cell’s life such as before vs. after it has differentiated.
2. **Regulating promoters and transcription.**  
   Assuming DNA is available for transcription (#1 above), cells regulate how many RNAs are transcribed & how fast by:
   1. controlling the assembly of the core promoter (TBP/TF-IID and related proteins);
   2. controlling specific transcription factors that can stimulate (enhancers) or inhibit (repressors) the activity of the core promoter; and
   3. turning up or down the rate at which RNA polymerase II moves along the DNA by modifying the C-terminal domain.
3. **Modifying RNA structure to regulate its half-life, and the final protein products.**As RNAs are transcribed:
   1. Splicing removes selected introns and creates alternatively spliced versions of mRNAs.
   2. The length of the polyA tail is adjusted to control mRNA activity and stability.
   3. The 5’ and 3’ UTRs bind other proteins that determine if or how long the mRNA will be used before being degraded.
4. **At the level of translation.**We did not talk much about it in class, but cells frequently adjust the rate of translation. Ways they can adjust translation rate include:
   1. Making miRNAs that block mRNA translation and tag it for destruction.
   2. Turning eIF4 on/off by chemical modification, which controls whether mRNAs are delivered to the primed 40S small ribosomal subunit.
   3. Turning eIF2 on/off by phosphorylation. eIF2 is needed to bring the first Met-tRNA into the small subunit.
   4. Turning GTP exchange on/off in eEF1. If this is turned off, the rate at which charged amino acyl-tRNAs are delivered to the A site in the large subunit slows down or stops.
   5. **All four of these mechanisms** are connected to signaling paths that tell the cell:
      1. if it has the resources it needs to make proteins, and
      2. if environmental conditions will let it make the proteins safely.
5. **Through post-translation modifications, complexes, compartmentalization, breakdown.**
   1. We have not talked about these much yet.
   2. In future topics we will see other ways that cells regulate the functions of proteins **after** they have been translated and folded.

***Why Do We CARE About Where and How Genes Are Regulated?***

First, it is essential to understand how they are controlled if we want to understand how cells work normally. Second, it helps us understand how cell functions break down in non-infectious, heritable diseases.

The first inheritable diseases for which we worked out mechanisms are caused by a single well-defined gene mutation. Examples include sickle cell anemia (caused by a single base mutation in beta globin) or hemophilia (caused by a mutation in the coding region of the gene encoding one of the clotting factors). However **most diseases with an inheritable predisposition are not easily mapped to one gene.**

We know that genetics contributes to cardiovascular disease, cancer, Alzheimer’s, and many other diseases, but we do not have clear causality where Mutation A causes Error B that leads to Outcome C. Instead, these disease or conditions are the result of **a network of faulty interactions** caused by:

* Inherited DNA mutations,
* Faulty epigenetic regulation,
* Improper transcription and translation of functional proteins, **and**
* External environmental factors

We are going to explore how maladaptive processes can **interact** with each other by looking at how they come together to disrupt normal brain development.

***Our Model System***

Autism spectrum disorders (ASD)are a family of neural developmental disorders defined by **four clinical features**:

1. Impaired ability to engage in social interactions or make social attachments.
2. Impaired speech and communication, especially in social situations.
3. Repetitive patterns of behavior or actions.
4. Restricted attention and interests.

There is a great deal of variation in the severity of these traits. This is partly why ASD is considered a family of related disorders.

Current evidence supports a model of ASD being caused by a c**ombination** of inheritable DNA mutations, faulty epigenetic and transcription controls, defective functional proteins, and environmental factors. We are going to leave out the environmental factors for this week and focus on the other elements that are occurring INSIDE of the neurons.

Your Goals and Tasks For This Case:

I selected 15 genes for which (as of August 2020) there is reliable, published evidence showing a **link** to autism spectrum disorders. I will provide you with a table that lists the standardized Gene ID in the human genome, full name, and a **general description of its function**. Some have additional information related to ASD or neural development.

NONE of these have a mutation that is sufficient by itself to be the single “cause” of autism. However, mutations or other changes in each of them have been shown to potentially CONTRIBUTE to autism.

Your goal is to identify the **smallest possible subset** of mutations or errors in these genes that, **acting together**, can produce the full set of abnormal behaviors and development changes we call ASD.

On **Wednesday** you will:

* Identify which of the distinctive traits of autism **each** gene might be causing or promoting.
* Decide whether each gene is involved in:
  + Epigenetic regulation,
  + Transcription control,
  + Translation control,
  + Neuron or brain development, or
  + General neuronal cell function.
* Propose a model for how autism develops that uses the smallest number of genes possible.

On **Friday**, your team have 5-8 minutes in class to explain your initial model.

* Please use an illustration or diagram showing what your model looks like.
* Hand-drawn sketches are fine; the point is to have something the other teams and I can LOOK at as you explain your thinking.
* **Send me your sketch in advance of class so I can project it.** You will have a chance to revise your model before you submit it on Saturday.

On **Saturday** what you will turn in is:

1. A finalized list of which genes are linked to epigenetic control, to transcription control, to translation and expression, and to end functions in the cell.
2. A short (1-2 paragraphs at most) description of your proposed model for how a minimum number of genes causes autism, and your reasoning behind it. Include a figure or diagram showing how the genes interact. (Basically, a revised version of what you presented on Friday.)

If you want more information about any gene, start with GeneCards (<https://www.genecards.org/>) using the Gene ID. For more on a specific disease mentioned in GeneCards or in another resource, use the MalaCards Disease Database (<https://www.malacards.org/>).)

Day 1 Data Handout:

|  |  |  |
| --- | --- | --- |
| **Clinical Features of ASD You Must Explain** |  | **Level of Control or Effect for Linked Genes** |
| Impaired ability to engage in social interactions or make social attachments. |  | Epigenetic regulation |
| Impaired speech and communication, especially in social situations. |  | Transcription control |
| Repetitive patterns of behavior or actions. |  | Translation control |
| Restricted attention and interests. |  | Neuron or brain development |
|  |  | General neuronal cell function |

***Table 1: Genes Linked to Autism***

The genes are listed in alphabetical order by the Gene ID.

|  |  |  |
| --- | --- | --- |
| **Gene ID** | **Gene Product Name, Information** | **Probably Acts At Level  of \_\_ & Why?** |
| ANKRD11 | Ankyrin repeat domain protein 11.   * Recruits histone deacetylases to chromatin, modifying histone acetylation and gene expression * Has a role in proliferation and development of cortical neuron precursors * Loss of this gene leads to intellectual disability. |  |
| ARID1B | AT-Rich Interaction Domain 1B.   * Part of SWI/SNF chromatin remodeling complex. * Regulates cell cycle activation, progression to mitosis. |  |
| CHD8 | Chromodomain Helicase DNA Binding Protein 8   * Unwinding of DNA in transcription, promotes cell proliferation, and regulates RNA synthesis. |  |
| FOXP2 | Forkhead box transcription factor P2.   * Binds directly to ~400 gene promoters. * Regulates a variety of genes. * Expressed in fetal and adult brain. * Mutation causes speech & language impairment, cognitive impairment, delayed motor development. |  |
| HIST1H1E | Histone 1.4.   * Binds to DNA between nucleosomes. * Mutants lack the lysines, arginines needed to hide negatively charged linker DNA, & for protein-protein interactions. |  |

|  |  |  |
| --- | --- | --- |
| IMMP2L | Inner Mitochondrial Membrane Peptidase Subunit 2.   * Transports proteins into mitochondria. * Controls one path to apoptosis. |  |
| KCTD13 | K+ channel tetrameric domain 13 protein   * Part of E3 ubiquitin-protein ligase complex. * E3 ligase sends RhoA to proteasome for destruction. Required for synapse formation. * Essential for behavioral self-regulation. |  |
| KDM5C | Lysine demethylase 5C   * Involved in the regulation of transcription and chromatin remodeling. * Mutations associated with X-linked cognitive disability. |  |
| MeCP2 | Methyl CpG binding protein 2   * Binds to DNA w/histone deacetylases (HDACs). * High concentrations in brain neurons. Associated w/ maturation of brain and forming new synapses. * Without MECP2, unrepaired DNA damage accumulates. * Mutation causes Rett syndrome (a type of ASD), with repetitive stereotyped behavior. * Children with non-Rett forms of ASD do not have MeCP2 mutations, but brain levels of MeCP2 are reduced. |  |
| OXTR | Oxytocin receptor   * Oxytocin/receptor are important in uterus during birth, & for forming parent-child attachment. * No alternative splice variants or sequence mutations in the coding region of this receptor have been found in ASD children. |  |
| RAY1/ST7 | * The gene is in a region on chromosome 7 known as ***autism-susceptibility locus***. * Sequencing of the region in autistic children did not find any specific mutations. * There is RNA transcribed from the gene, but no known proteins are translated from the RNA. |  |
| RELN | Reelin   * Secreted protein regulates neural cell migration in developing brains. * After birth it regulates synapse remodeling. |  |
| SHANK3 | Shank protein 3.   * Scaffold protein in dendrites that connects neurotransmitter receptors and ion channels to the cytoskeleton and holds them in place. * Needed for synapse formation on dendrites. |  |

|  |  |  |
| --- | --- | --- |
| UBE3A | Ubiquitin protein ligase E3A.   * Breakdown of RhoA is required for synaptic transmission. * Ubiquitin ligase E3A is part of the complex that sends RhoA to the proteasome for destruction. * Mutation causes Angelman Syndrome: severe motor and intellectual retardation, muscle weakness, seizures, loss of speech, facial defects. |  |
| ZNF778 | Zinc Finger Protein 778.  (Zinc finger proteins are zinc-containing proteins that form specific DNA binding sites.   * Required for normal development of cognitive skills. * Mutations cause cognitive impairment. |  |

### Day 1 Summary

By the end of class **today** you should have:

* Identified which of the distinctive traits of autism **each** gene might be causing or promoting.
* Decided whether each gene is involved in:
  + Epigenetic regulation,
  + Transcription control,
  + Translation control,
  + Neuron or brain development, or
  + General neuronal cell function.
* Started sketching out a model for how autism develops that uses the **smallest number of genes possible**.

Day 2 Instructions:

Today, your team will have 5-8 minutes in class to explain your initial model.

* Please use an illustration or diagram showing what your model looks like.
* Hand-drawn sketches are fine; the point is to have something the other teams and I can LOOK at as you explain your thinking.
* **Send me your sketch via team folders so I can project it.** You will have a chance to revise your model before you submit it.

Tomorrow you will turn in:

1. Your finalized list of which genes are linked to epigenetic control, to transcription control, to translation and expression, and to end functions in the cell.
2. A short (1-2 paragraphs at most) description of your proposed model for how a minimum number of genes causes autism, and your reasoning behind it. Include a figure or diagram showing how the genes interact. (Basically, a revised version of what you presented on Friday.)

Case and Instructor Notes:

### Background

How genes regulate neural development has been extremely difficult to work out. Current evidence suggests >4100 human genes contribute to development of the brain and nervous system. It has long been thought that genes controlling neural development are arranged as a series of “**network modules**.” Dozens or even hundreds of genes work as a functional group, and one cell type can have dozens of different, active network modules. Data suggest at least 2983 of the >4100 genes involved in human neural development are part of at least one network module. Some of these genes likely belong to more than one network. The other genes may not be controlled as part of a network, or we simply have not found their network yet.

It also is well-established that one protein or regulatory RNA can have many different effects in the nervous system at different developmental stages and in different cell types. This suggests that neural cells have epigenetic programming which regulates how they respond.

Epigenetic mechanisms regulate chromatin structure and gene expression without altering the DNA sequence. They play an important role in fine-tuning development-related genes and epigenetic dysregulation can cause neurodevelopmental disorders, including ASD. Micro- and lncRNA are additional epigenetic regulators that can control the expression of many genes simultaneously at the level of post-transcription by blocking protein synthesis or mRNA degradation. Abnormal regulatory RNA synthesis can lead to neurodevelopmental disorders.

### Case Data Structure

The scenario and candidate genes in this case come from 4 articles:

* Krol, et al. 2018. Epigenetic modification of the oxytocin receptor gene is associated with emotion processing in the infant brain. <https://doi.org/10.1016/j.dcn.2019.100648>
* Loke, et al. 2015. The role of epigenetic change in autism spectrum disorders. <http://dx.doi.org/10.3389/fneur.2015.00107>
* Wu, et al. 2020. Recent Progress on Relevant microRNAs in Autism Spectrum Disorders. <http://dx.doi.org/10.3390/ijms21165904>
* Yoon et al. 2020. Genetic and Epigenetic Etiology Underlying Autism Spectrum Disorder. <http://dx.doi.org/10.3390/jcm9040966>

Several genes have long been thought to be involved in autism. A cluster of these genes lie in Chromosome 7q22-q33. Several of these genes are also involved in other neurological disorders.

* **Forkhead box P2 (FOXP2):** transcription factor for lung, CV, intestinal, and neural tissues. Point mutation is linked to language deficits
* **RAY1/ST7:** contains a long noncoding RNA called ST7 overlapping antisense 1-4 (ST7OT1-4) that may regulate the gene in which it resides.
* **IMP2 inner mitochondrial membrane protease-like (IMMP2L):** in IMMP2L knockdown mice, deficiency induced behavioral effects that were gene-dose and sex dependent.
* **Reelin (RELN):** trinucleotide GGC repeat in the 5' untranslated region alters expression. Having more copies correlates with vulnerability to ASD. Family-based association analysis for 218 Caucasian families found RELN was an important potential contributor to autism.

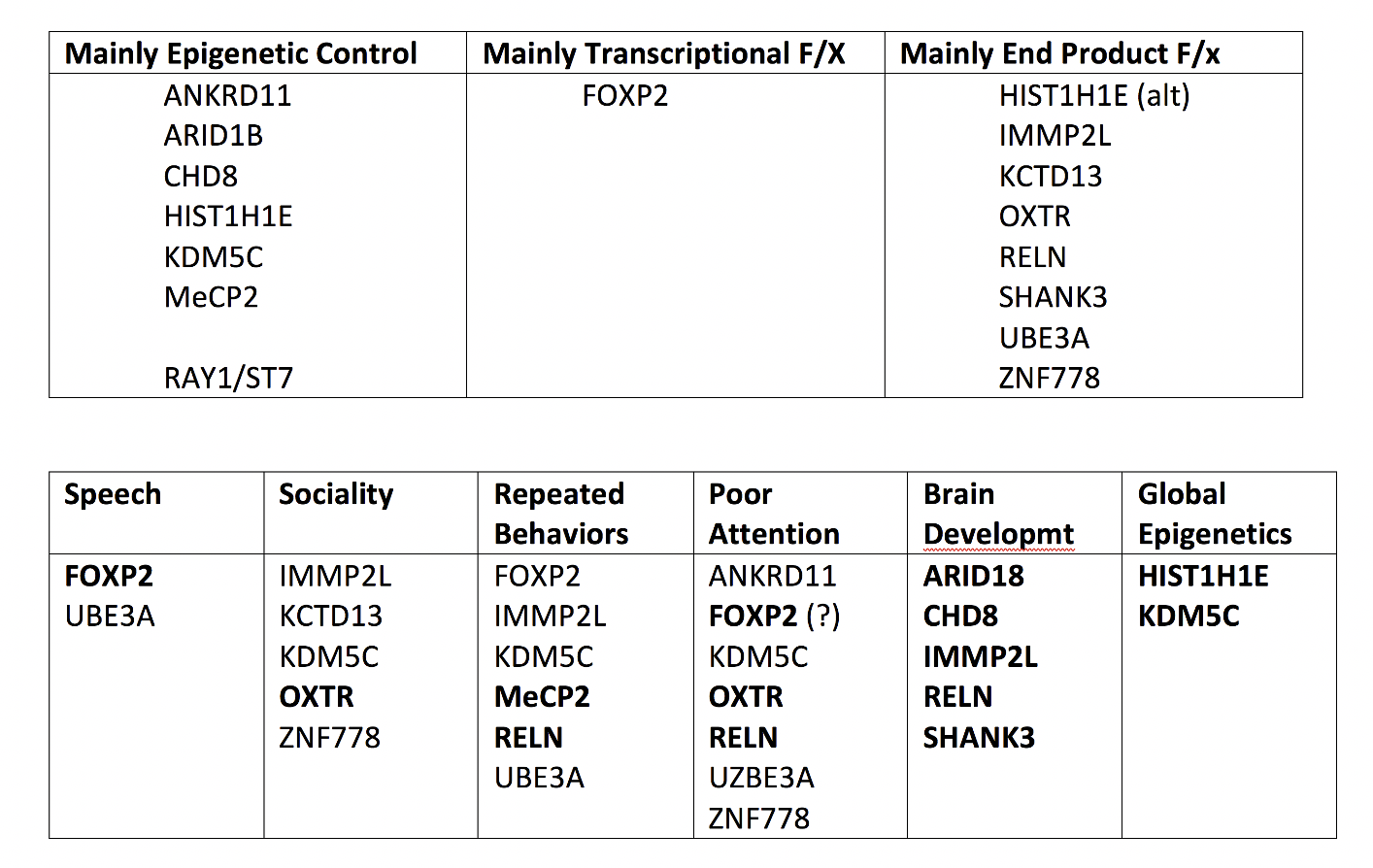
Other genes associated with autism lying outside of the Chr. 7q cluster include:

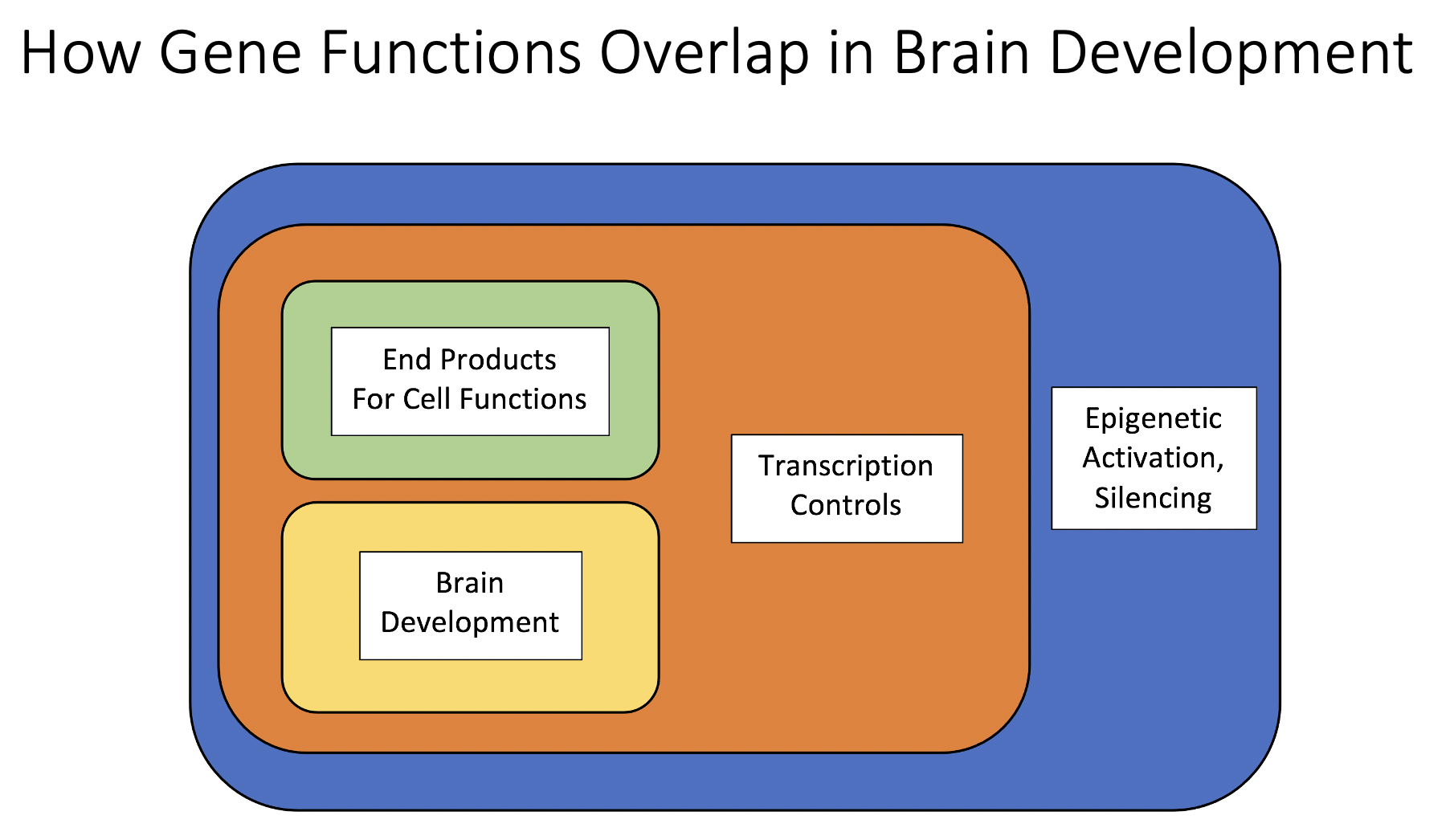
* **Ubiquitin-protein ligase E3A (UBE3A).** Loss of imprinting of one copy, and production of antisense RNA that binds to UBE3A and mRNA prevents translation.
* **Gamma aminobutyric acid receptor subunit beta 3 (GABRB3)**
* **MeCP2:** Encodes a methyl binding protein that binds to methylated region of DNA and silences the gene. Has a role in synaptic development and long-term synaptic plasticity by recruiting co-repressors, and controlling chromatin looping.

Many students come to this case thinking eukaryotic gene regulation is essentially the same as prokaryotes (i.e., the operon model). One goal is to show students that eukaryotic gene regulation is fundamentally different. So when choosing the list of candidate genes I intentionally favored those associated with epigenetic regulation, and genes with multiple associated disease phenotypes.

### Typical Results

The tables below show how the genes can be classified based on functions, and the most likely points where they affect neural development in ASD.





### Class Management

Both cases before this one in my class have step-by-step instructions for completing them. Starting with this case I intentionally give students fewer instructions, for two reasons.

1. I want them to start working through more complex problems within their teams. I let teams struggle with the assignment for a few minutes, and if they are not able to start, I give them one clue that sends them in the right direction. It helps that there is no one answer to the puzzle; their logic and reasoning are more important.
2. This case comes shortly before their first exam, where they must think through complex questions and formulate responses on their own. Practicing something similar with their team greatly reduces exam anxiety.

##### Goals of the In-Class Presentations

There is no one correct set of genes or arrangement that students need to find. I focus my questions on the logic that they used to connect their minimum number of candidates. About 1 in 4 groups will make an error that invalidates their model, but they have an opportunity to correct or modify their models before submitting them for a grade.

### Supplemental Activities and Case Extensions

* To make the case more challenging, give students the original articles and ask them to identify the genes they should evaluate.
* Part 2 (next section) has not been tested with students, but would be appropriate for an advanced class. Students are asked to develop a network model for ASD and use it to generate experimentally testable hypotheses.

Part 2: Gene Regulation in Autism 2

### Background: What Can We Learn From Network Modules?

Network modules make it easier to identify which genes contribute to abnormal neural development. If we find that Gene A is associated with a particular condition, other genes that are in network modules with Gene A are more likely to be involved in that abnormality too. If we have mapped out the network module, we can predict which genes may be more or less central for producing the abnormal conditions.

Human speech is an example of how defective chromatin structure, transcription, and translation come together to produce abnormal neural development. When young children are exposed to people talking, they develop the ability to understand words, and convert their own thoughts into spoken words. Most develop this skill without being taught.

Two conditions illustrate that speech development is genetically controlled. First, there is an extended family in England with a unique inheritable defect in the FOXP2 gene. This is a transcription factor associated with neural development of speech and language centers in the brain. Affected members of the family have normal intelligence and ability to understand language but cannot form syllables or speak understandably.

Second, children with childhood apraxia of speech (CAS) cannot learn to speak, though they have normal intelligence and speech-associated anatomy. Most CAS children have a defect in one of the genes linked to CAS listed below; ALL of these genes are part of one network module that has high expression during early and mid-fetal brain development.

(Don’t focus on details; look at what the genes linked to CAS are DOING generally.)

##### **Epigenetic Control Genes Linked to CAS**

|  |  |
| --- | --- |
| **Gene ID in Humans** | **Gene Name and Function** |
| BCL11A | **BAF Chromatin Remodeling Complex Subunit 11A**  DNA binding protein that regulates gene expression via **chromatin remodeling.** |
| SETD1A, WDR5 | **SET domain histone methyltransferase 1; WD repeat domain protein 5**  Both proteins are part of histone-3 lysine-4 methylation complex. Defects linked to schizophrenia, intellectual defects, speech/language delays |
| CHD3 | **Chromodomain-helicase-DNA-binding protein 3**  Part of Mi-2/NuRD histone deacetylase complex. |
| ANKRD12 | **Ankyrin repeats-containing cofactor 12**  Inhibits transcriptional activity of steroid/nuclear receptors by attracting histone deacetylases. |

##### **Transcription Regulation Genes Linked to CAS**

|  |  |
| --- | --- |
| **Gene ID in Humans** | **Gene Name and Function** |
| ATF2 | **Activating transcription factor 2**  Binds to cAMP-responsive element (CRE), an 8-base sequence, and activates cAMP-responsive genes.  **Histone acetyltransferase (HAT) for histones H2B and H4 near CRE elements**. |

##### **Functional Proteins With Translation Errors (Mutations) Linked to CAS**

|  |  |
| --- | --- |
| **Gene ID in Humans** | **Gene Name and Function** |
| ERC1 | **E3 Ubiquitin Ligase 1**  Guanine nucleotide exchange factor (GEF) that controls membrane **vesicle trafficking, neurotransmitter release.** |
| STX5 | **Syntaxin 5**  Member of t-SNARE family of proteins required for **synaptic vesicle docking and fusion to membrane.** Has a critical role in autophagy. |
| EPPK1 | **Epiplakin 1**  Cytoskeletal linker protein. Controls reorganization of intermediate filaments in response to cell stress, injury. |
| HCN2 | **Hyperpolarization Activated Cyclic Nucleotide Gated K/Na Channel 2**  Part of automatic pacemaker currents in heart and neurons. Linked to sense of taste for sweets. |

Based just on this list of network module genes and their functions, we can make some generalizations about neural development in CAS.

1. Defects in gene regulation contribute to CAS as much as mutations in genes encoding functional proteins.
2. General histone modifications are an important part of CAS. These include H3K4 methylation [activation of regions], general deacetylation [deactivation] of large regions, and acetylation [activation] of specific cAMP-responsive genes.
3. Defects in movement and release of neurotransmitter vesicles are probably involved in CAS development.

From this very brief analysis, we now have three general ideas we can use to develop specific testable hypotheses.

### Your Task

1. Go back to the list of genes linked to autism in Part 1 of this case. Following the same logic as shown above, what generalizations could we make about neural development in ASD?
2. Assume for a minute that we have a mouse model of ASD that matches how humans develop it. Design an experiment to test one of your generalizations. What is your working hypothesis?