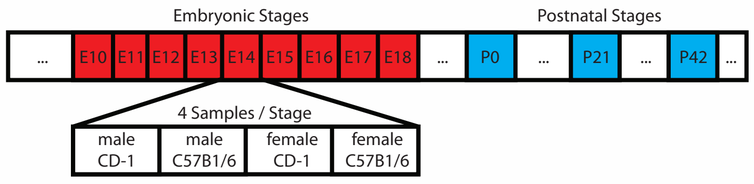
Developmental Expression Pattern of SET Domain Genes

Daryanaz Dargahi, Tyler Funnell, Calvin Lefebvre, Shing Hei Zhan

Background

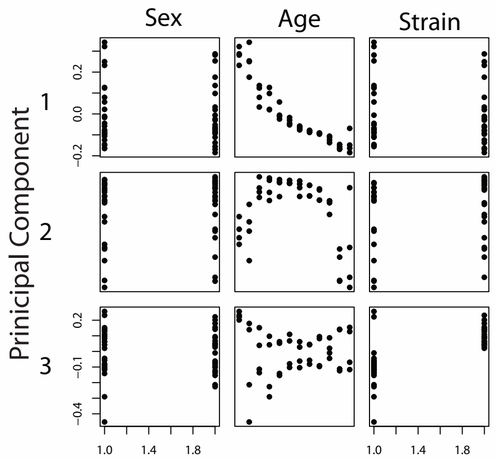
- Paediatric overgrowth syndromes are characterized by tall stature, dysmorphic features, and intellectual disability.
- The pathogenesis of these diseases is thought to originate from abnormalities in the hypothalamus—a growth hormone-secreting gland (e.g., Sotos syndrome [1,5]).
- Several overgrowth syndromes are caused by germ-line mutations in SET domain genes (e.g., the Sotos syndrome gene NSD1 [3] and the Weaver syndrome gene EZH2 [2]).
- The highly conserved SET (Su(var)3-9 and 'Enhancer of zeste') domain is involved in the epigenetic control of gene transcription.
- A recently published genome-wide expression atlas of the normal mouse hypothalamus [6] provides a platform for formulating hypotheses concerning the genetic regulation of mammalian embryonic development.
- We utilized this data set to:
  1. explore the role of SET domain genes in guiding normal hypothalamus development
  2. speculate how germ-line perturbation in SET domain genes may lead to overgrowth syndromes by adversely impacting hypothalamus development.

Microarray Data



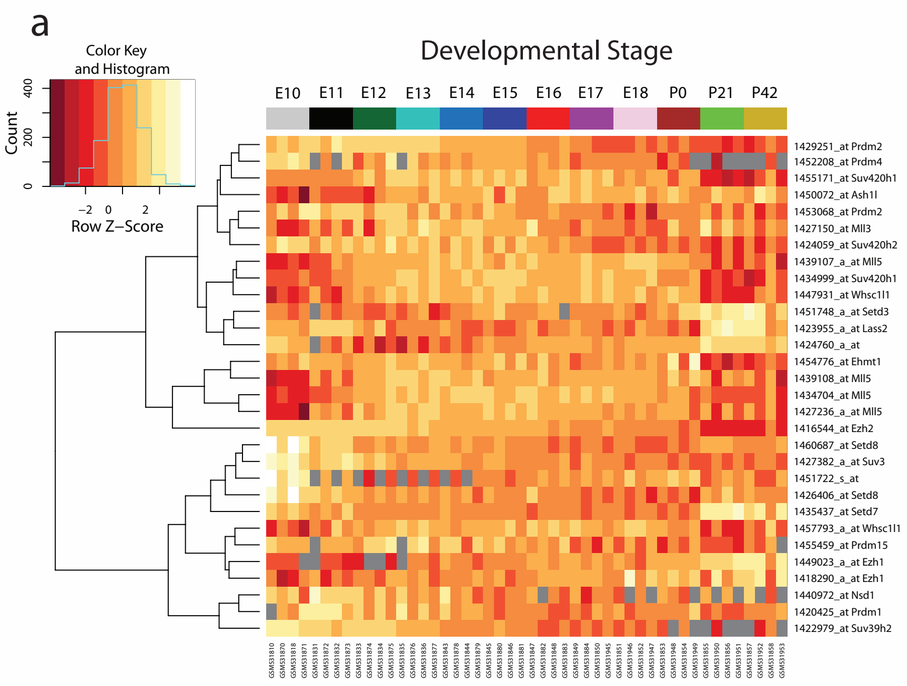
Data Pre-Processing Procedure:

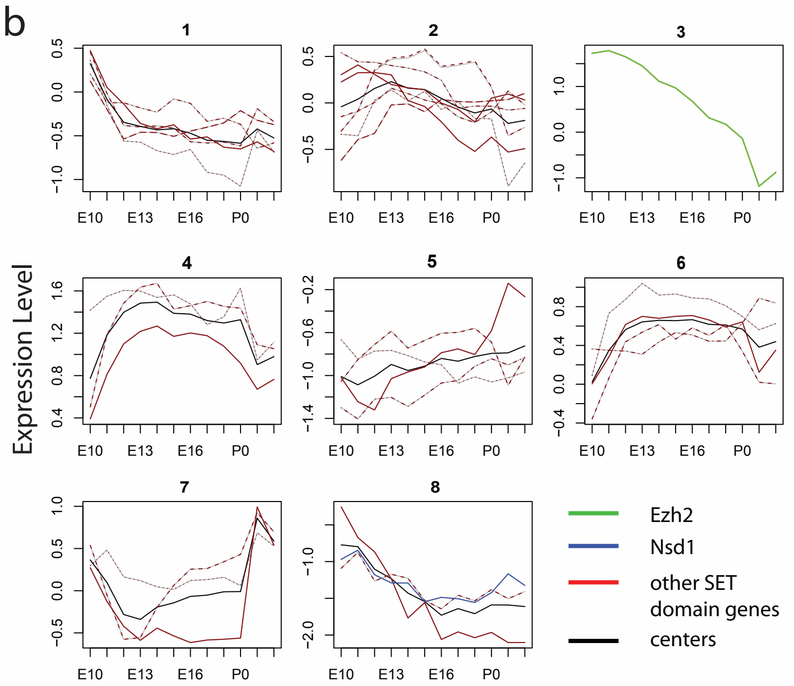
1. Remove probes with absent or marginal absolute calls according to detection p-value
2. Remove probes having missing values in more than 10 samples
3. Log transform
4. Remove probes with fewer than 2 values of at least log2(3)
5. Quantile normalize
6. Check for batch effect (by age, strain, developmental stage, and the first three principal components)



The first 3 principal components are not associated with either sex or strain. The first principal component, however, is correlated with age.

Clustering of Expression Profiles





Similarities among the expression profiles of differentially expressed SET domain genes (detected using limma [7]) hint at co-regulatory mechanisms involved in normal hypothalamus development (**a**). Further clustering of differentially expressed SET domain genes (using the fuzzy *c*-means method implemented in Mfuzz [4]) revealed a notable diversity of temporal expression patterns among SET domain genes (**b**). Nsd1, homolog of the human Sotos syndrome gene, has an expression profile similar to those of Prdm1 and Suv39h2, both previously implicated in tumourigenesis.

Summary of SET Domain Genes

Gene Name	Ensemble ID	Unigene ID	Disease Associated with Human Homolog	Cluster
Ezh2	ENSMUSG00000029687	Mm.246688	Weaver Syndrome, myeloid disorders	3
Ezh1	ENSMUSG00000006920	Mm.5027	Breast cancer	5
Nsd1	ENSMUSG00000021488	Mm.168965	Sotos, Weaver, and Beckwith-Wiedemann syndromes, liver cancer	8
Ehmt1	ENSMUSG00000036893	Mm.24176	Kleefstra syndrome	4
Prdm1	ENSMUSG00000038151	Mm.4800	B cell lymphoma, melanoma, stomach cancers	8
Prdm2		Mm.332020	lymphoma; liver, colorectal, stomach, and breast cancers	2
Prdm4		Mm.25307	pancreatic, ovarian, and stomach cancers	1
Prdm15	ENSMUSG00000014039	Mm.328741	leukemia	5
Setd3	ENSMUSG00000016599	Mm.159185		6
Setd7	ENSMUSG00000037111	Mm.192111		7
Setd8		Mm.137966		1
Mll3		Mm.332268	myeloid disorders	2
Mll5		Mm.205190	leukemia	4, 6
Suv39h1	ENSMUSG00000039231	Mm.9244	sarcoma, B cell lymphoma	1
Suv39h2	ENSMUSG00000026646	Mm.128273	head/neck cancer, B cell lymphoma	8
Suv420h1	ENSMUSG00000045098	Mm.278578		2
Suv420h2	ENSMUSG00000045098	Mm.278578		2
Lass2	ENSMUSG00000015714	Mm.181009		7
Ash1l	ENSMUSG00000028053	Mm.130752	breast cancer	2
Whsc1l1		Mm.217337	Wolf-Hirschhorn syndrome	5, 6

Remarks

- 22 (out of 42 covered in the pre-processed data set) SET domain genes are temporally differentially expressed in normal mouse hypothalamus.
- Groups of SET domain genes exhibit distinct expression profiles, suggesting that SET domain genes may be regulating different genetic network modules.
- We propose that the 22 differentially expressed SET domain genes to be prioritized for future experimental work to further understand the mechanistic role of SET domain genes in the development of hypothalamus and in the pathogenesis of paediatric overgrowth syndromes.
- We also propose to conduct similar exploratory analyses on other organs involved in growth and development (e.g., pituitary gland) to further elucidate the global role of SET domain genes in mammalian development.

References

1. Cole (1990) *J Med Genet* **27**: 571-576.
2. Gibson et al. (2012) *Am J Hum Genet* **90**: 110–8.
3. Kurotaki (2002) *Nat Genet* **30**: 365-6.
4. Matthias and Futschik (2010) Mfuzz version 2.12.0.
5. Ranke and Bierich (1983) *Eur J Pediatrics* **140**: 109-11
6. Shimogori et al. (2010) *Nat Neurosci* **13**: 767–75.
7. Smyth (2004) *Stat Appl Genet Mol Biol* **3**: article 3.