Supplemental information

De novo variants implicate chromatin modification,

transcriptional regulation, and retinoic acid

signaling in syndromic craniosynostosis

Andrew T. Timberlake, Stephen McGee, Garrett Allington, Emre Kiziltug, Erin M. Wolfe, Amy L. Stiegler, Titus J. Boggon, May Sanyoura, Michelle Morrow, Tara L. Wenger, Erica M. Fernandes, Oana Caluseriu, John A. Persing, Sheng Chih Jin, Richard P. Lifton, Kristopher T. Kahle, and Paul Kruszka

Table of Contents:

Figure S1. Venn diagram demonstrating overlap between syndromic CS risk genes and genes implicated in autism, intellectual disability, and other NDDs.

Disease	Genes in DisGeNET	CS Candidate Genes Overlapping with DisGeNET	P-Value
Neurodevelopmental Disorders	535	31	✦ 6.55×10^{-20}
Autism Spectrum Disorders	1071	30	● 5.83×10^{-11}
Intellectual Disability	2165	60	▲ 4.16×10^{-22}

Figure S1. Venn diagram demonstrating overlap between syndromic CS risk genes and genes implicated in autism, intellectual disability, and other NDDs. Overlap between genes implicated in CS and each phenotype was significant by Fisher's exact test (Methods).

Figure S2. High pLI genes with multiple damaging *de novo* variants in syndromic CS probands.

Figure S2. High pLI genes with multiple damaging *de novo* variants in syndromic CS probands. Quantile-quantile plot of log transformed observed vs expected P-values for genes with damaging *de novo* variants in syndromic CS probands (Poisson distribution, see Methods). Thirteen genes surpassing the threshold for genome-wide significance after Bonferroni correction for high pLI genes $(0.05/3{,}063 = 1.63 \times 10^{-5})$ are labeled.

Figure S3. Domain annotations for CHD3, CDK13, MED13L

A. CHD3
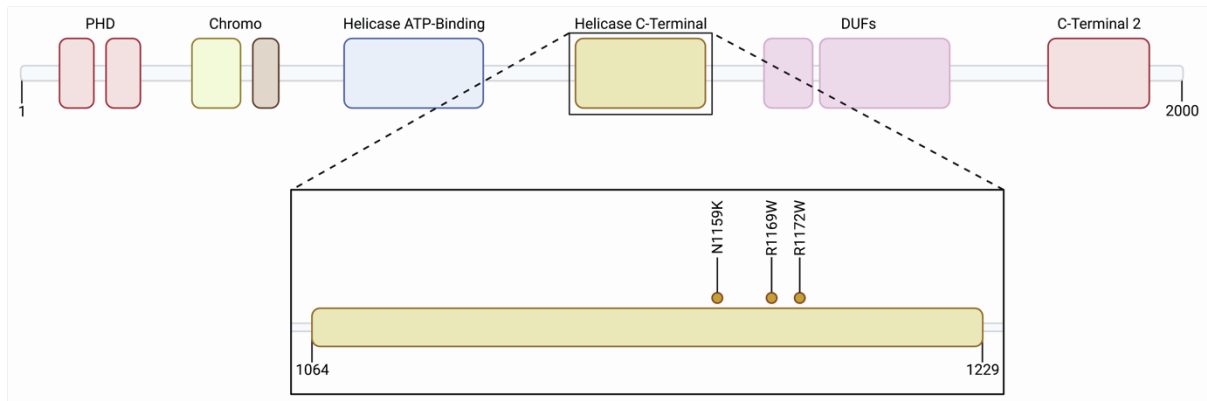
B. CDK13
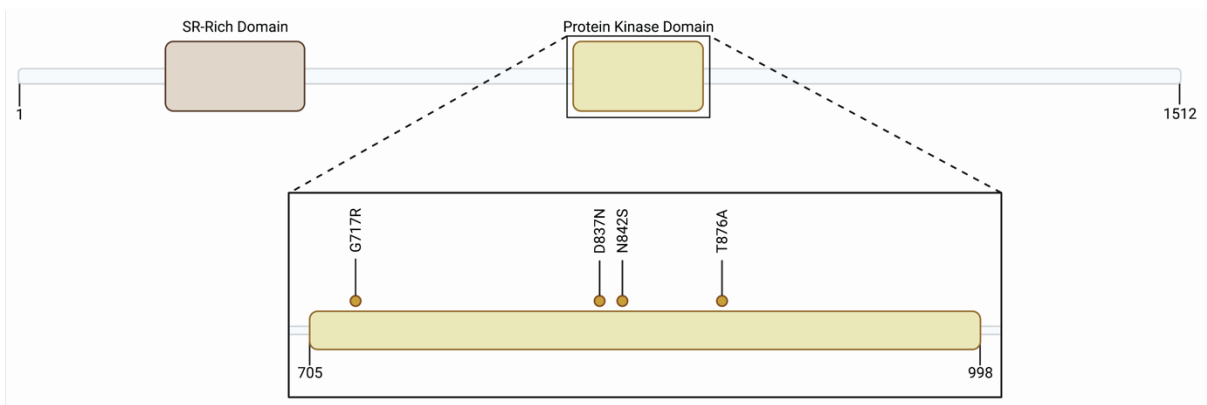
C. MED13L
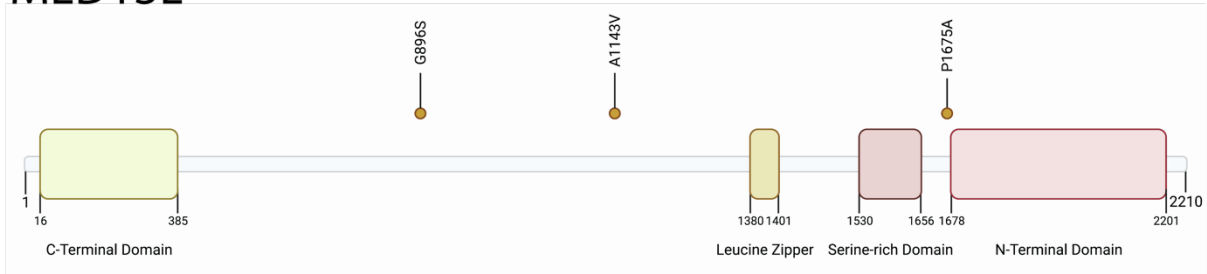
Figure S3. Domain annotations for CHD3, CDK13, MED13L. Domain annotations appear for genes in which damaging *de novo* variants identified in syndromic CS probands were uniquely missense variants. **A.** In CHD3, missense variants are all in the helicase C-terminal domain. **B.** In CDK13, all 4 variants appear within the kinase domain. **C**. In MED13L, variants are scattered throughout the protein, supporting haploinsufficiency as a possible mechanism of contribution to CS risk.

Figure S4. Pathway analysis of genes in modules enriched in CS.
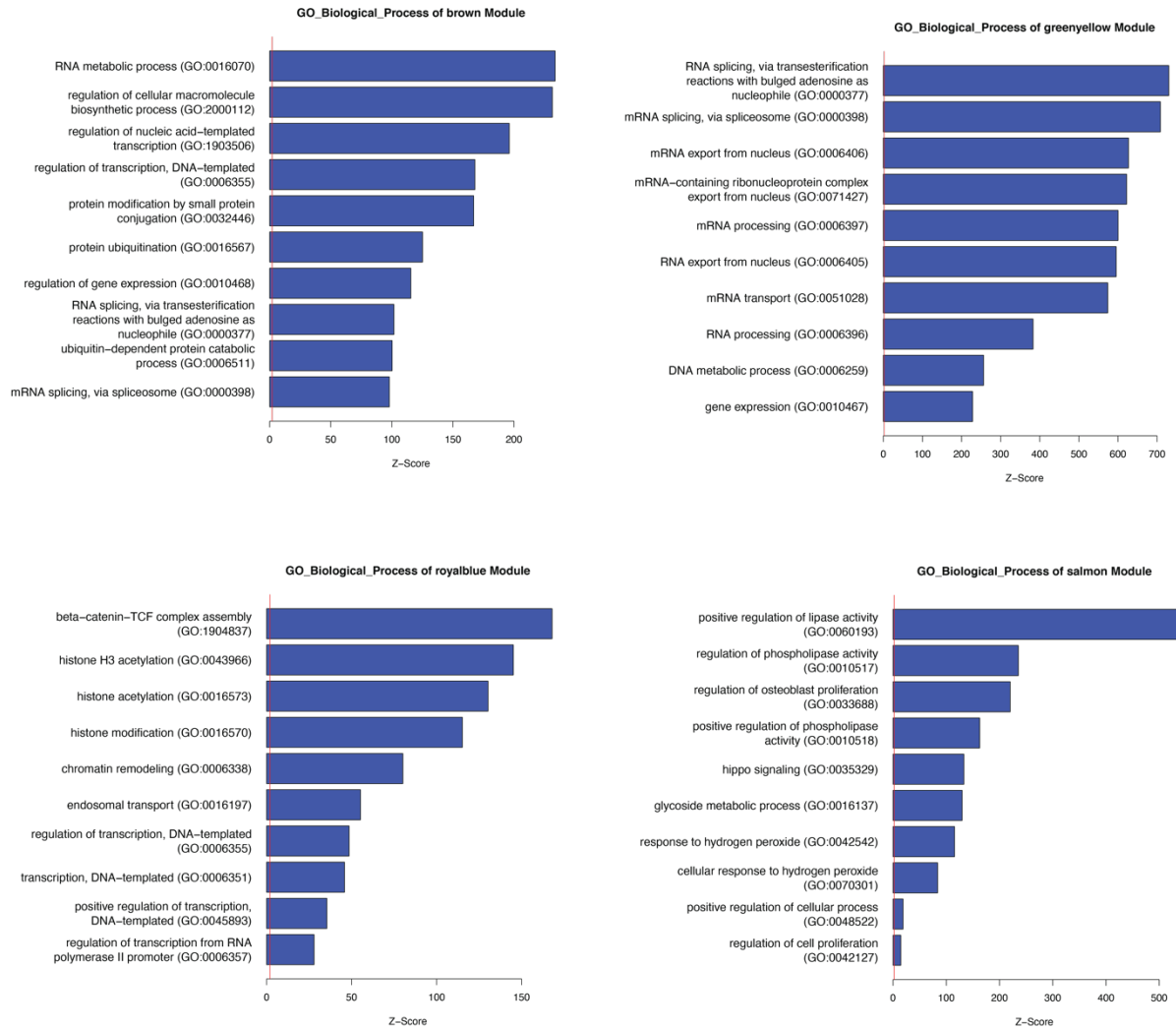
Figure S4. Pathway analysis of genes in modules enriched in CS. Pathway analysis of genes encompassed by each of the four transcriptional networks enriched for genes mutated in syndromic CS. The brown module genes are implicated in transcriptional regulation, the greenyellow module genes are implicated in RNA splicing, the royalblue module genes are implicated in histone modification, and the salmon module genes are involved in osteoblast differentiation.

Figure S5. Temporal expression of coexpression networks implicated in syndromic CS.
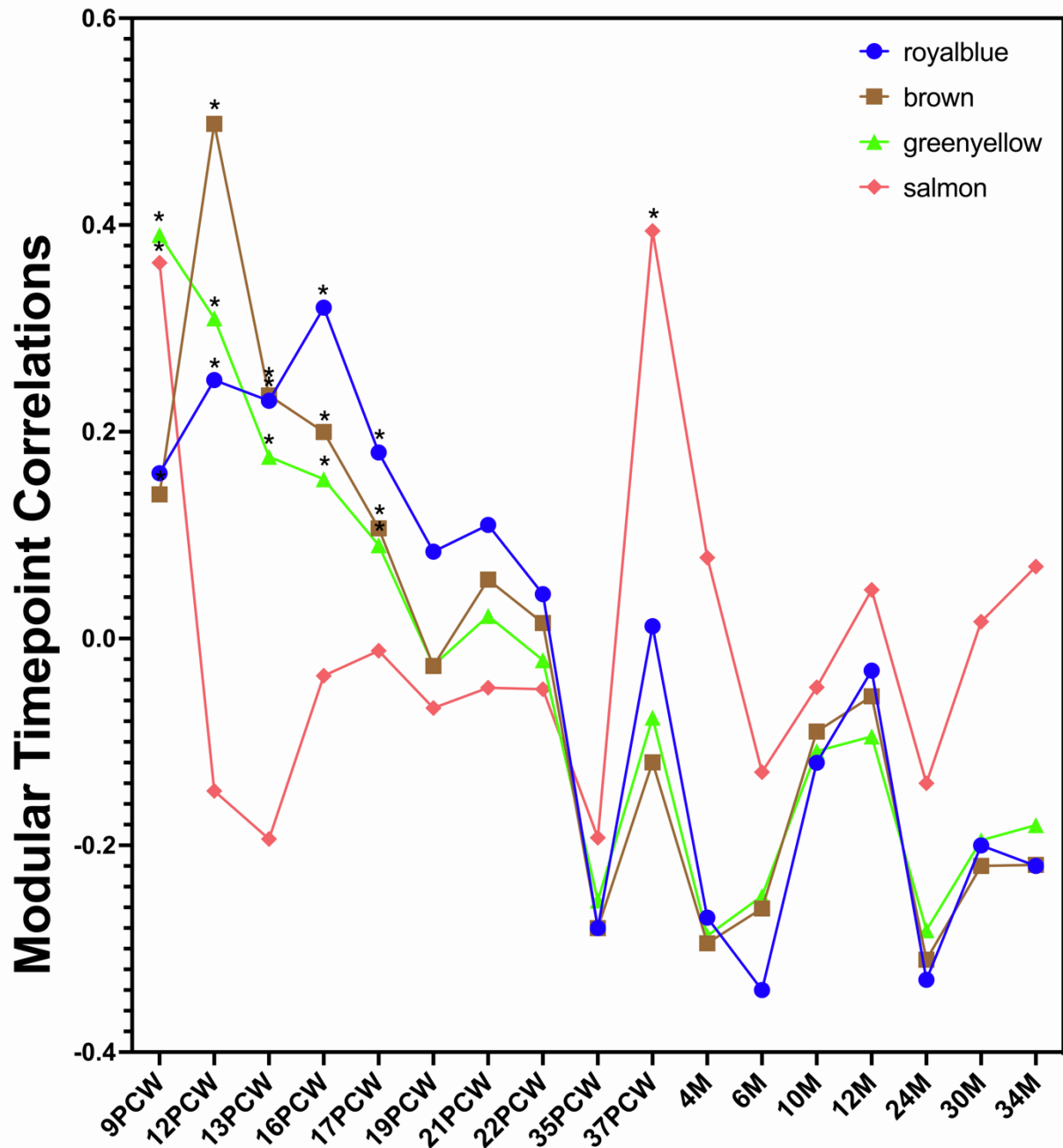
Figure S5. Temporal expression of coexpression networks implicated in syndromic CS. Temporal expression of the brown, greenyellow, royalblue, and salmon modules are plotted, with periods with significantly enriched expression indicated with asterisks (Methods). Implicated modules are highly expressed during the mid-gestational period during which the skull ossifies and sutures form, namely post-conception weeks 9-17.

Figure S6. UMAP plot of scRNA-seq data demonstrating 32 distinct clusters.
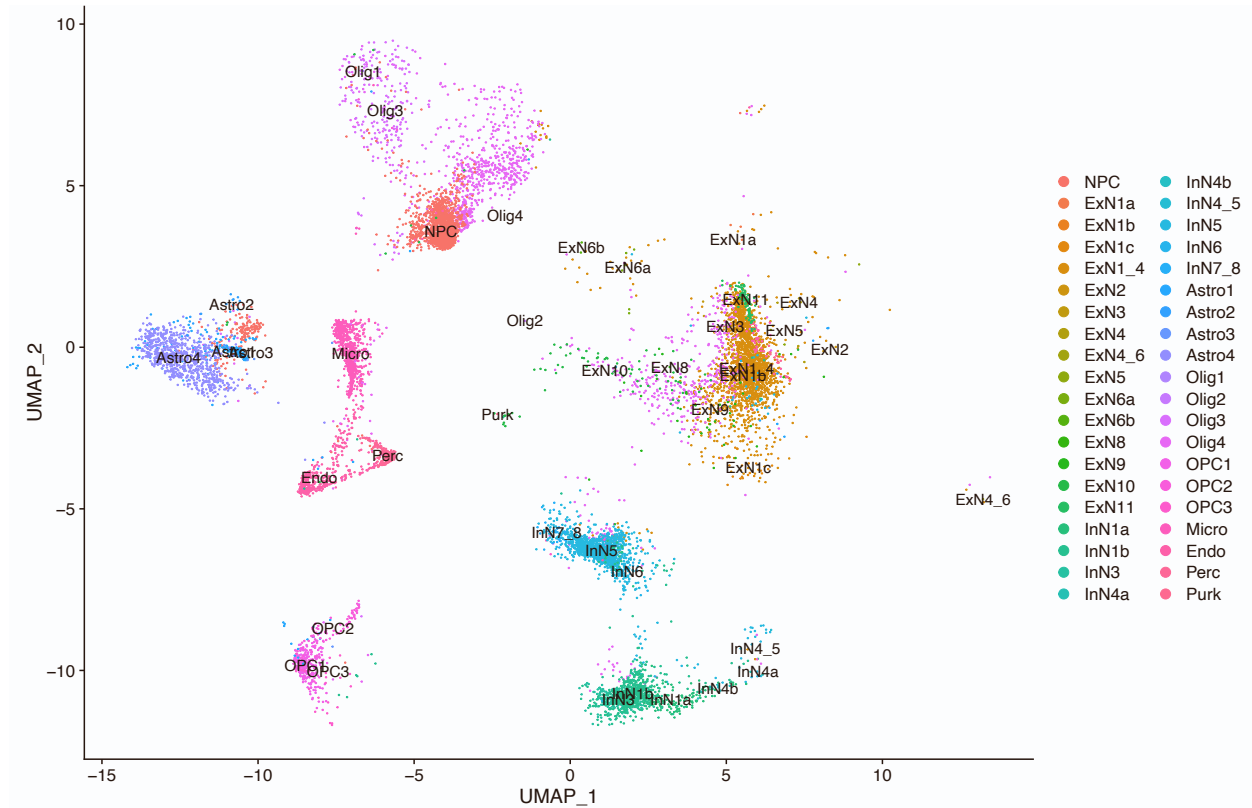
Figure S6. UMAP plot of scRNA-seq data demonstrating 32 distinct clusters. Clusters identified using over 40,000 cells were used to identify specific neural cell populations, which were assessed for enrichment in CS risk genes (Methods). NPC, neural progenitor cells; ExN, excitatory neurons; InN, inhibitory neurons; Olig, oligodendrocytes; OPC, oligodendrocyte progenitor cells; Micro, microglia; Astro, astrocytes; Endo, endothelial cells; Perc, pericytes; Purk, purkinjee cells.

Figure S7. Estimation of the number of syndromic craniosynostosis risk genes by simulation.
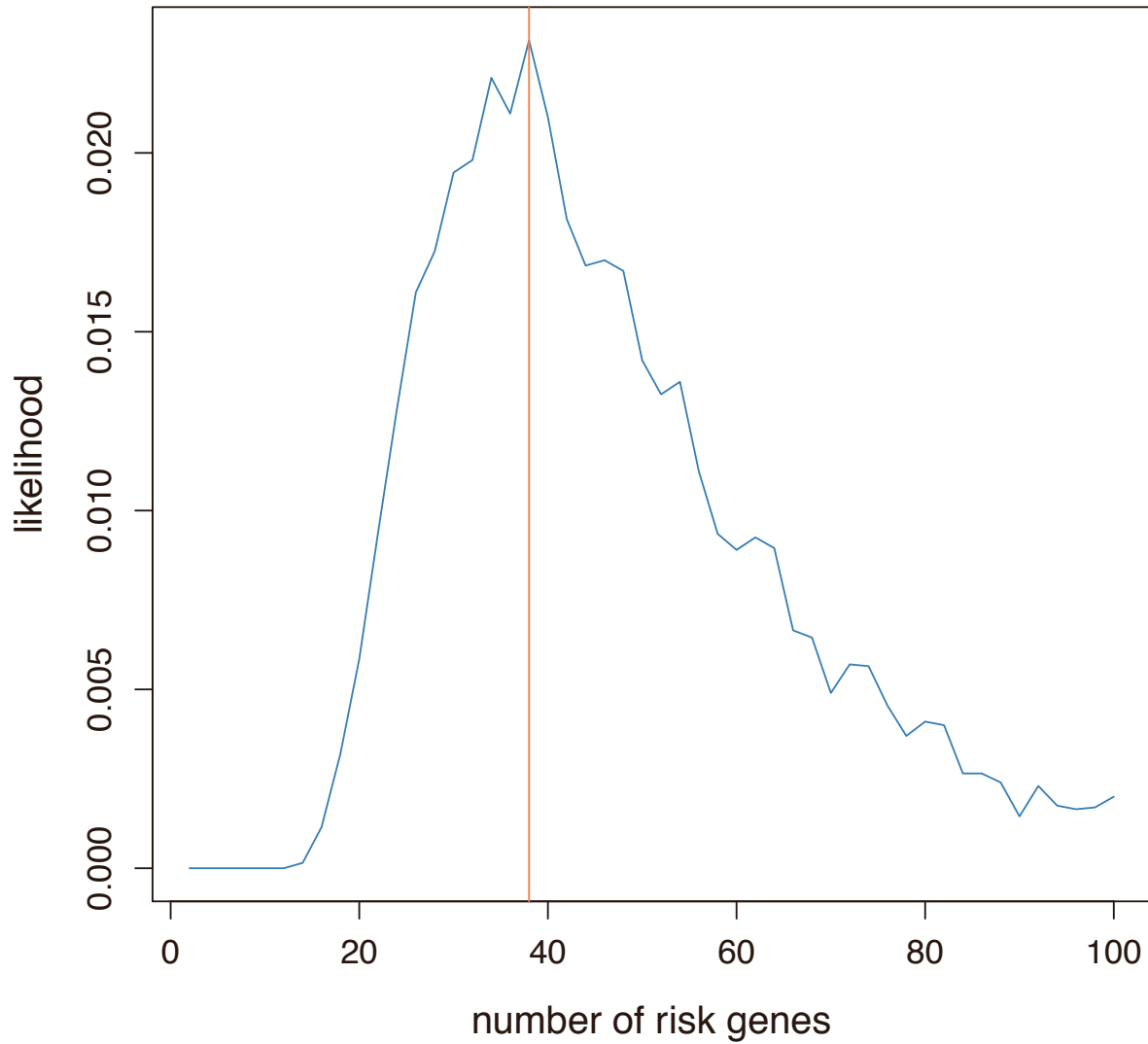
Figure S7. Estimation of the number of syndromic craniosynostosis risk genes by simulation.
The likelihood of the model under the assumption of there being 0-100 risk genes is plotted, with maximum likelihood observed at 38 genes (Methods).

Table S1. Curated List of Neurodevelopmental Disorder, Autism, and Craniosynostosis Risk Genes (provided as Excel File Table S1)

Table S2. Neural Crest Cell Gene Expression Rank (provided as Excel File Table S2)

Table S3. Known syndromic CS genes

Genes	
ADAMTSL4	IL11RA
ALPL	IRX5
ALX4	JAG1
ASXL1	KAT6A
ATR	KMT2D
CDC45	KRAS
COLEC11	LMX1B
CTSK	LRP5
CYP26B1	MEGF8
EFNA4	MSX2
EFNB1	PHEX
ERF	POR
ESCO2	RAB23
FAM20C	RECQL4
FBN1	RUNX2
FGFR1	SCARF2
FGFR2	SH3PXD2B
FGFR3	SKI
FLNA	SPECC1L
GLI3	STAT3
GNAS	TCF12
GNPTAB	TGFBR1
GPC3	TGFBR2
HUWE1	TMCO1
IDS	TWIST1
IDUA	WDR35
IFT122	ZEB2
IHH	ZIC1

Known syndromic CS genes, derived from Twigg et al[2]. Probands were screened for variants in these genes, and those with identified variants were removed from further assessment.

Table S4. *De novo* variants in known genes in syndromic CS probands

Gene	Impact	GnomAD Frequency	pLI
EFNB1	p.Arg109Gly	0	0.93
EFNB1	p.Gln214*	0	0.93
EFNB1	p.Arg66*	0	0.93
ERF	p.Ser297fs*9	0	0.99
FAM20C	p.Gly280Arg	4.9×10^{-6}	0.31
FGFR2	p.Glu566Gly	0	1
FGFR2	p.Try105Cys	4.0×10^{-6}	1
FGFR2	p.Tyr340Cys	0	1
FGFR2	p.Ser347Cys	0	1
FGFR2	c.G1032A; p.=	0	1
FGFR3	p.Pro250Arg	7.4×10^{-6}	0
FGFR3	p.Ala391Glu	0	0
FGFR3	p.Ala391Glu	0	0
GNAS	p.Asp799Asn	0	0.68
HUWE1	p.Arg110Trp	0	1
HUWE1	p.Arg110Trp	0	1
HUWE1	p.Arg110Gln	0	1
JAG1	p.Ile819Leufs*2	0	1
KAT6A	p.Arg1024*	0	1
KAT6A	p.Ser1880*	0	1
KMT2D	c.15922-1G>A	0	1
KMT2D	p.Ile3420Val	0	1
KMT2D	p.Asn4027His	0	1
KRAS	p.Gly60Ser	4.0×10^{-6}	0
SH3PXD2B	p.Ile633Val	0	0
SKI	p.Asn128Ser	0	1
TGFBR2	p.Arg553Cys	0	0.13
ZEB2	p.Leu349*	0	1
ZIC1	p.Asp348Asn	0	0.94

Variants identified in known syndromic CS genes in 555 trios with sporadic CS. The impact of each variant is provided at the protein level for missense, frameshift, and nonsense variants, and at the DNA level for splice variants. Complete annotations for each variant are found in Table S6. pLI= probability of LOF intolerance. Allele frequencies and pLI scores were obtained from GnomAD v2.1.1. In the case of *FGFR2*, *FGFR3*, and *KRAS* variants found once in GnomAD, each is also annotated as 'pathogenic' by ClinVar. Variants in known genes explained 5.2% of cases studied.

Table S5. Phenotypes of patients with mutations in novel CS risk loci (provided as Excel File Table S5)

Table S6. All DNVs identified in 526 probands with unsolved syndromic CS (provided as Excel File Table S6)

Table S7. GOrilla pathway analysis of all genes harboring damaging DNVs (provided as Excel File Table S7)

Table S8. Burden of *de novo* variants in 1,789 autism control trios

	Observed		Expected		Enrichment	P
	n	Rate	n	Rate		
High pLI Genes (>0.9) (n=3,063)						
Total	453	0.25	491.3	0.27	0.92	0.96
Synonymous	113	0.06	137.1	0.077	0.83	0.98
Total missense	305	0.17	309.2	0.18	0.99	0.60
D-mis	61	0.034	66.2	0.037	0.92	0.76
LoF	35	0.020	45.0	0.025	0.78	0.95
Damaging	96	0.054	111.2	0.062	0.86	0.93
High pLI OMIM Genes (n=643)						
Total	111	0.062	120.3	0.067	0.92	0.81
Synonymous	25	0.014	33.7	0.019	0.74	0.95
Total missense	80	0.044	75.9	0.042	1.05	0.33
D-mis	20	0.011	25.9	0.014	0.77	0.90
LoF	6	0.0034	10.8	0.0060	0.56	0.96
Damaging	26	0.015	36.7	0.021	0.71	0.97
Chromatin Modifiers and Remodelers (n=614)						
Total	77	0.043	87.4	0.049	0.88	0.88
Synonymous	17	0.0095	23.7	0.013	0.72	0.94
Total missense	54	0.031	55.2	0.031	0.98	0.58
D-mis	16	0.0089	13.0	0.0073	1.23	0.23
LoF	6	0.0034	8.5	0.0048	0.71	0.85
Damaging	22	0.012	21.5	0.012	1.02	0.48

n, number of *de novo* variants in 1,789 control trios; Rate, number of *de novo* variants per subject; Damaging missense called by MetaSVM (D-mis); Loss of function denotes premature termination, frameshift, splice site variant, startloss, or stoploss variants. P-values represent the upper tail of the Poisson probability density function.

Table S9. Enrichment in chromatin genes in 498 trios with syndromic CS after exclusion of 38 probands with variants in 13 genes identified via *de novo* variant burden at the individual gene level

Class	Observed #	Observed #/subject	Expected #	Expected #/subject	Enrichment	P-value
All variants	40	0.08	24.7	0.05	1.62	*0.003*
Synonymous	6	0.01	6.76	0.01	0.88	0.67
T-mis	12	0.02	11.9	0.02	1.00	0.53
D-mis	13	0.03	3.65	0.01	3.56	*1.14x10^{-4}*
Loss of function (LOF)	9	0.02	2.38	0.005	3.78	*8.13x10^{-4}*
Damaging	22	0.04	6.03	0.01	3.65	*4.23x10^{-7}*

D-mis as called by Meta-SVM. Damaging refers to D-mis and LOF alleles. P-values represent the upper tail of the Poisson distribution; significant P-values appear in italics.

Table S10. *De novo* variants in high pLI OMIM genes in syndromic CS probands

Gene	Impact	GnomAD Frequency	pLI
AHDC1	p.Gln1156*	0	1
CACNA1E	p.le1451Met	0	1
COL11A1	c.3468+1G>A	0	1
COL2A1	p.Arg904Cys	0	1
COL3A1	p.Arg1358Gln	1.2×10^{-5}*VUS	1
COL4A1	p.Gly888Arg	0	1
CTNNA1	p.Val374_375insSerTrpLysMetLys	0	0.97
CTNND1	p.Arg458*	0	1
FBN2	p.Cys1156Phe	0	1
HCN4	p.Leu438Val	0	1
KCND3	p.Val401Met	0	0.99
MAST1	p.Arg496His	0	1
MFN2	p.Arg95Thr	0	0.99
NAA15	p.Leu89Pro	0	1
NF1	p.Thr2135Ile	0	0.90
PDHA1	p.Arg340Cys	0	0.99
PRKG1	p.Pro537fs	0	1
SCN1A	p.Val846Ile	0	1
SLC20A2	p.Pro568Leu	1.2×10^{-5}*P	0.97
SOX4	p.Gly94Asp	0	0.93
TRIO	p.Gln1427Arg	0	1
USP9X	c.7219-2A>G	0	1
ANKRD11	p.Ala1751Ser	0	1
BICD2	p.Lys734Asn	0	0.98
CLTC	p.Pro890Leu	0	1
CSNK2A1	p.Arg47Gln	0	0.99
FBXO11	p.Pro109Leu	0	1
FOXF1	p.Asn219Ser	0	0.96
HNRNPK	p.Glu85Lys	0	1
HNRNPK	p.Asp190Ala	0	1
HTT	p.Thr341Ser	0	1
HTT	p.Met1623Val	0	1
KANSL1	p.Arg942Trp	0	1
MCM6	p.Cys158Tyr	0	0.98
KIF11	p.Leu517Val	0	1
MED12	p.Val1964Leu	0	1
MTOR	p.Arg2018Pro	0	1
NOTCH1	p.Ala16Thr	0	1
PKD1	p.Ser322Leu	1.7×10^{-5}	1
PPP2CA	p.Ser30Thr	0	0.99
ROBO2	p.Ser545Thr	0	1
SPOP	p.Thr25Ala	0	1
YAP1	p.Ala351Thr	0	1
ZBTB20	p.Gly63Arg	5.6×10^{-5}	0.97
ZNF462	p.Ala1171Thr	1.1×10^{-5}	1

Variants identified in known high pLI OMIM genes in 526 trios with sporadic syndromic CS not identified at the individual gene level. The impact of each variant is provided at the protein level for missense, frameshift, and nonsense variants, and at the DNA level for splice variants. pLI= probability of LOF intolerance. Allele frequencies and pLI scores were obtained from GnomAD v2.1.1. Gene names in bold represent LOF or D-mis alleles, while those not in bold represent missense variants not predicted to be damaging by MetaSVM, as these also demonstrated significant enrichment in probands. *VUS represents a variant in ClinVar described as a variant of unknown significance, and *P represents a ClinVar variant designated pathogenic.

Table S11. Enrichment of LOF intolerant genes with high expression in cranial neural crest cells in 526 trios with syndromic CS

Class	Observed		Expected		Enrichment	P-value
	#	#/subject	#	#/subject		
All variants	148	0.28	82.7	0.16	1.79	*6.75x10-11*
Synonymous	22	0.042	23.0	0.044	0.96	0.61
T-mis	53	0.10	41.3	0.079	1.27	*0.04*
D-mis	39	0.07	10.8	0.02	3.62	*2.55x10-11*
Loss of function (LOF)	34	0.06	7.63	0.015	4.46	*2.11x10-12*
Damaging	73	0.14	18.4	0.035	3.97	*6.51x10-22*

D-mis as called by Meta-SVM. Damaging refers to D-mis and LOF alleles. P-values represent the upper tail of the Poisson distribution; significant P-values appear in italics. Genes in the top 25% of CNCC expression with pLI>0.9 (n=1,593) were included for gene set enrichment analysis.

Supplementary Note 1: Probability of observing two identical *de novo HUWE1* (p.Arg110Trp) or *RARA* (p.Gly289Arg) variants.

Calculating this probability is analogous to the "birthday paradox"; i.e., the chance that in a set of "n" randomly chosen people, at least one pair will have the same birthday.In this scenario, we let R denote the total number of two identical DNVs in the cohort; R_j denotes the number of two identical DNVs in the jth tri-nucleotide category (e.g., TCC -> TTC; based on flanking base context, there are 4*4*4*3=192 categories in total); M denotes the total number of DNVs; M_j is the number of DNVs in the jth tri-nucleotide category.

Additionally, denote the number of potential DNVs in the jth tri-nucleotide category as L_j; per-base variant probability of the jth category is μ_j. The quantity of interest is then:

$$\mathbb{P}(R > 0 \mid M = m)$$

$$\approx \sum_{j=1}^{192} \mathbb{P}(R_j > 0 \mid M = m)$$

$$= \sum_{j=1}^{192} \sum_{k=2}^{m} \mathbb{P}(R_j > 0, M_j = k \mid M = m)$$

$$= \sum_{j=1}^{192} \sum_{k=2}^{m} \mathbb{P}(R_j > 0 \mid M_j = k) \times \mathbb{P}(M_j = k \mid M = m)$$

To calculate $\mathbb{P}(R_j > 0 \mid M = m)$:

$$\mathbb{P}(M_j = k \mid M = m) = \binom{m}{k} \times [\mathbb{E}(M_j \mid M = 1)]^k \times [1 - \mathbb{E}(M_j \mid M = 1)]^{m-k}$$

$$= \binom{m}{k} \times (\frac{\mu_j L_j}{\sum_{i=1}^{192} \mu_i L_i})^k \times (1 - \frac{\mu_j L_j}{\sum_{i=1}^{192} \mu_i L_i})^{m-k}$$

We use the Poisson tail probability to estimate $\mathbb{P}(R_j > 0 \mid M_j = k)$:

$$\mathbb{P}(R_j > 0 \mid M_j = k) \approx 1 - \exp\left(-\binom{k}{2}/L_j^2\right)$$

Taken together, we have

$$\mathbb{P}(R > 0 \mid M = m) \approx \sum_{j=1}^{192} \sum_{k=2}^{m} \left(1 - \exp\left(-\binom{k}{2}/L_j^2\right)\right) \times \binom{m}{k} \times \left(\frac{\mu_j L_j}{\sum_{i=1}^{192} \mu_i L_i}\right)^k \times \left(1 - \frac{\mu_j L_j}{\sum_{i=1}^{192} \mu_i L_i}\right)^{m-k}$$

We use this equation to estimate the probability of having two identical DNVs given a total of 609 DNVs ($P = 6.5 \times 10^{-3}$).

Supplementary Note 2: Phenomic Analyses

Compressing the HPO tree for dimensionality reduction

There are currently >16,000 disease-related terms in the Human Phenotype Ontology (HPO). To perform our phenomics analyses, we needed to reduce the number of features to a more manageable number by first group similar phenotype terms together. Terms in the ontology that are more similar based on Hybrid Relative Specificity Similarity (HRSS) are grouped together using a bottom-up approach to hierarchical clustering. Initially, all HPO terms begin as separate clusters. Then, cluster pairs are joined together with the most similar going first. Using this process, the optimal number of clusters (phenotype groups) can be adjusted based on a desired specificity. In this analysis, we compressed the HPO tree down to 250 phenotype groups. Each patient is then represented by a 250-length vector, where the ith number represents the count of HPO terms that this patient exhibits within the ith phenotype group.

Running UMAP/HDBSCAN

Using these vectors, we can then perform Uniform Manifold Approximation and Projection (UMAP) for dimensionality reduction, enabling us to visualize our patients in two-dimensional space. In this two-dimensional space, we then perform Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) to assign cluster IDs to each of the patients. Of the 526 patients, we were able to successfully cluster 498 samples together across 6 separate phenotype clusters (shown below). The resulting 98 samples which did not cluster with the other groups were omitted from the remainder of the visualizations.
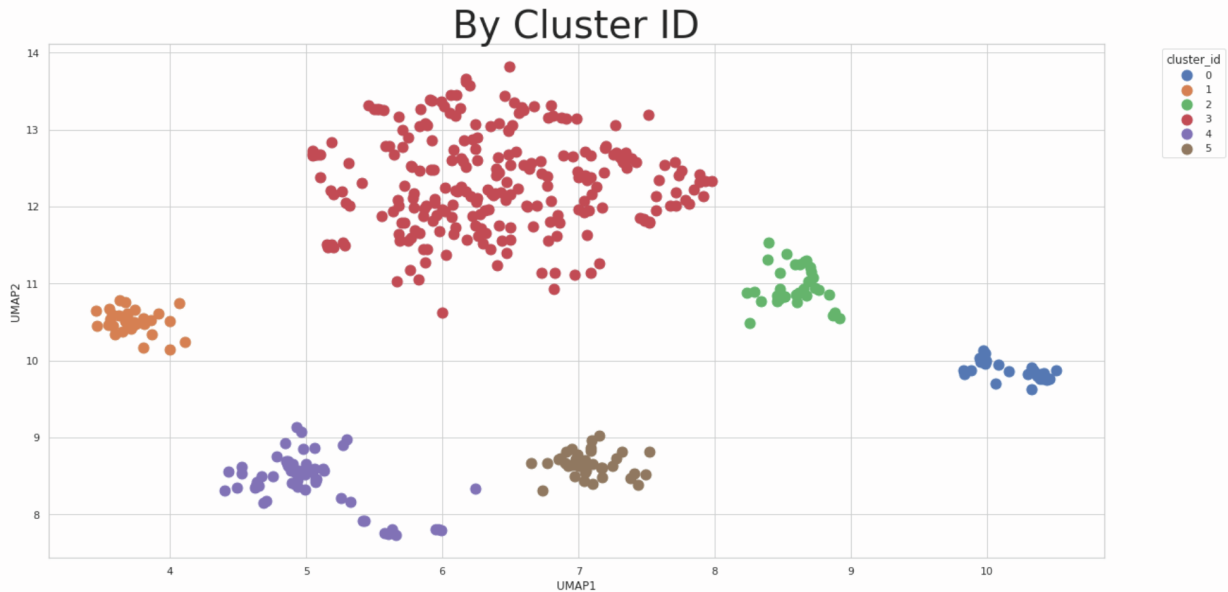

Basic cluster stats

In looking at each cluster, we show the proportion of females, mean and standard deviation of age. As a reminder, cluster -1 pertains to patients which were not assigned to a cluster using our HDBSCAN algorithm.

| Cluster	N	prop_female	mean_Age	std_age
-1	98	0.45	4.98	6.09
0	30	0.40	3.54	3.45
1	33	0.39	4.93	4.94
2	33	0.36	6.75	8.13
3	240	0.39	5.57	7.24
4	56	0.32	6.51	6.64
5	36	0.25	7.51	8.20

Phenotype group frequencies per cluster

To further characterize each cluster, we generated a series of waterfall plots showing the frequencies of phenotype groups that each set of patients exhibit. Each term in the waterfall plot represents the most-used term from the phenotype group in our in-house dataset. For instance, in group 1, 100% of patients exhibit a phenotype term that resides in the same phenotype group (and is therefore most similar to) "Scoliosis". This allows us to quickly evaluate the defining feature for several groups (e.g. – Group 4 appears to predominantly exhibit "seizure"-related phenotypes).
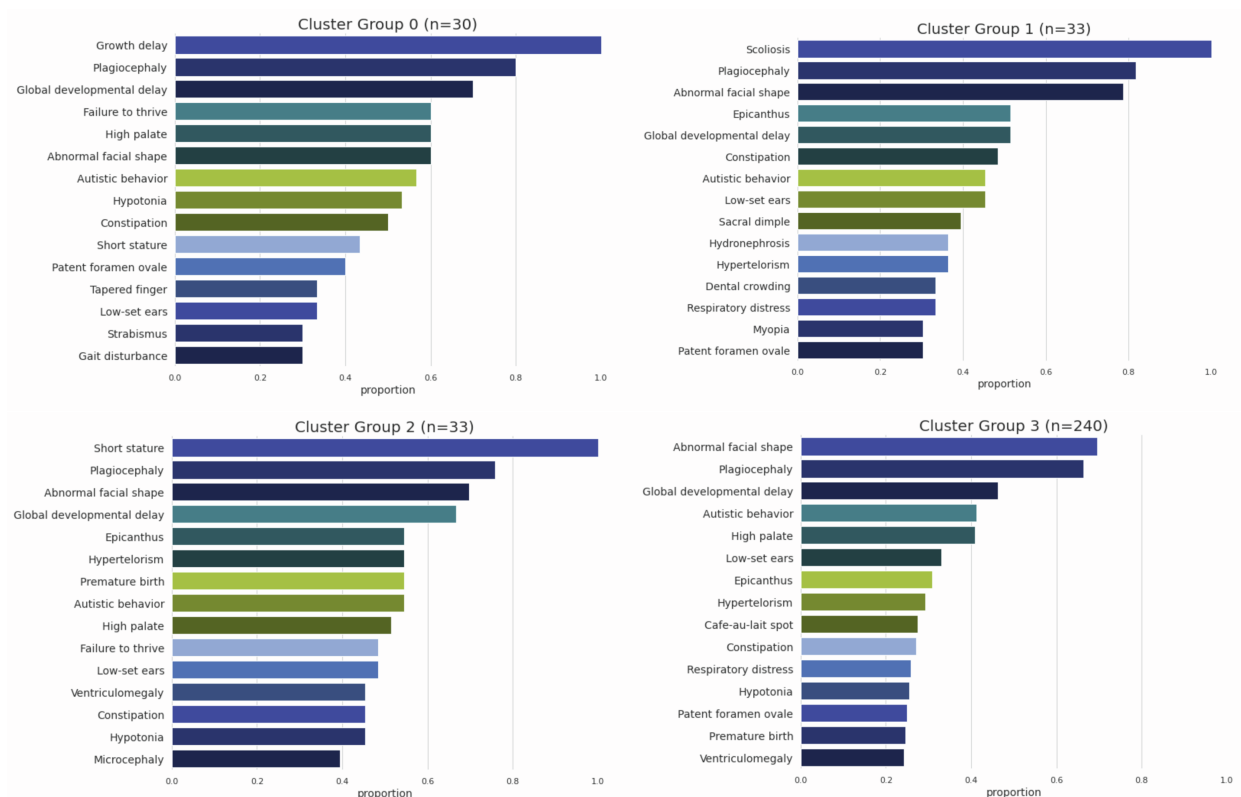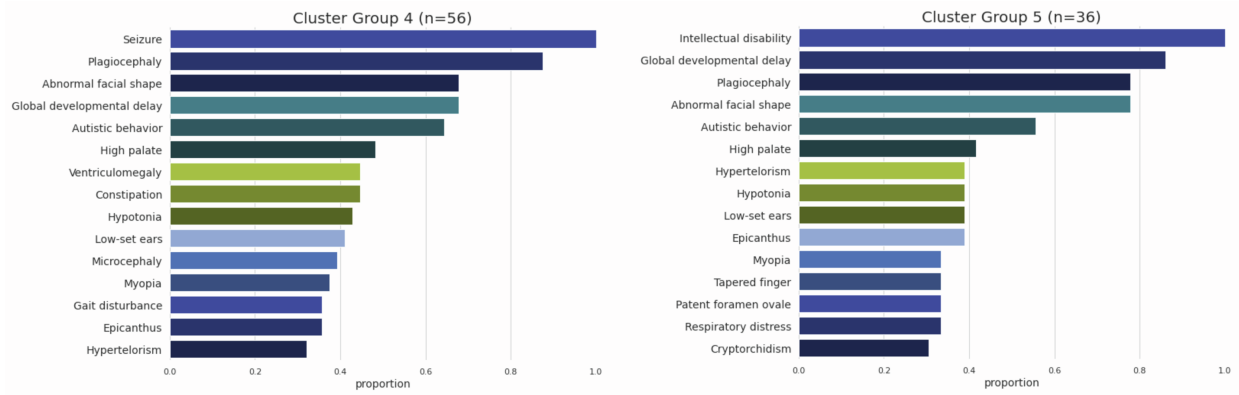
Can we better characterize each cluster?

In looking at the waterfall plots above, we see that "plagiocephaly" appears in every cluster, indicating that it is not a particularly differentiating factor. To define the most differentiating phenotypes within each phenotype cluster, we relied on a statistical technique, Term Frequency-Inverse Document Frequency (TF-IDF). This method is commonly used in document analysis to define the most important words within a document, while down weighting terms that occur in many documents. In our case, TF-IDF would essentially ignore "plagiocephaly" as a differentiating phenotype given that it exists in all clusters. The image below shows the normalized TF-IDF values for all terms in each cluster. The top terms per cluster have been annotated.
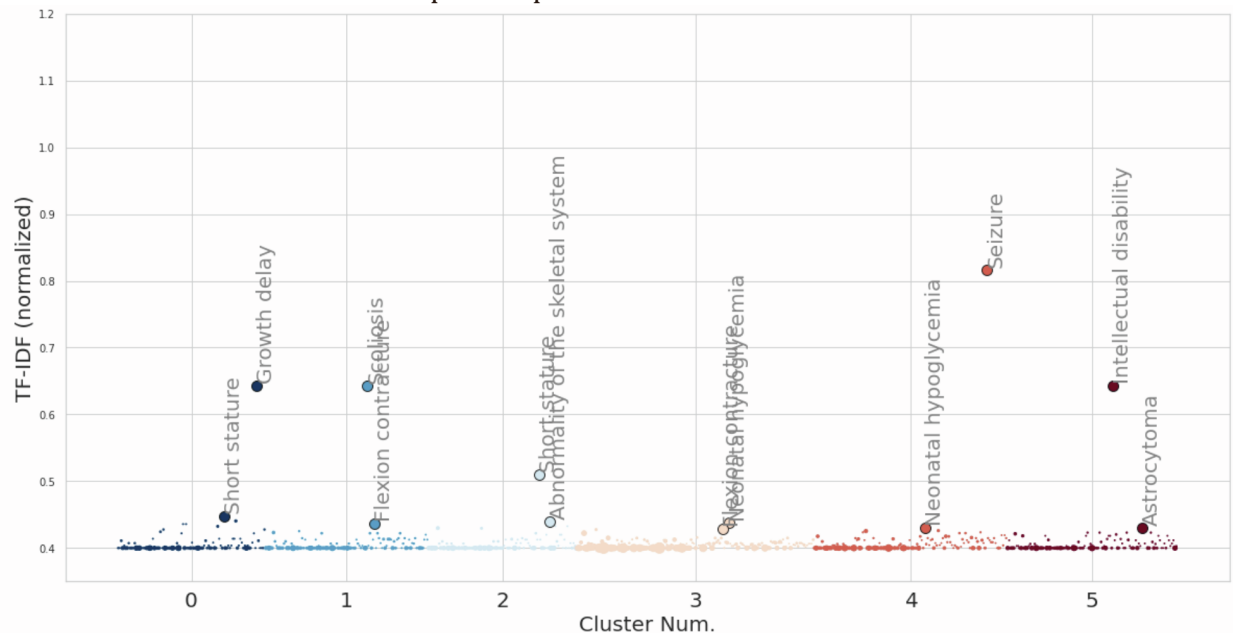
While we were able to identify the above clusters, these patients did not share similar genotypes in most all cases, thus we can't make any conclusions about specific genotype/phenotype relationships with our current number of samples. Instead, these results suggest that mutations in these pleiotropic genes affect various organ systems as expected, but a larger number of patients with mutations in each will be necessary to fully characterize the phenotypic impact of mutations in each.

Supplementary Note 3: Case Reports

Clinical synopses of probands with a recurrent *de novo* variant in *RARA*

Proband 1:
Patient is 21-month-old male
Birth: full term, 8 lbs 7 oz, 20 inches long, head size "larger"
Prenatal: no concerns. Negative noninvasive prenatal testing.
Newborn: Difficulty feeding due to ankyloglossia, poor weight gain. Noted to have bilateral rocker bottom feet (found to be vertical tali) and abnormal head shape (scaphocephaly). Diagnosed with sagittal craniosynostosis.

· CT head: Fusion of the anterior sagittal suture. The remainder of the sutures are patent.

Diagnoses:
- Bilateral vertical tali with calcaneovalgus deformities
- Partial sagittal craniosynostosis (HC at 17 months 48 cm, 78th%)
- Poor weight gain
- Failure to thrive (weight at 19 months 7.79 kg, 0.05 percentile, z score -3.27)
- Short stature (length at 19 months was 74 cm, 0.02%, z score -3.49)
- Hypotonia
- Unilateral cryptorchidism with hernia
- Ankyloglossia

Infancy: poor weight gain necessitating NG tube feedings. Hypotonia and motor delay. Hospitalized at 3m of age due to poor weight gain. Sagittal suturectomy performed in early infancy. NG tube was feeding was stopped at 16 months of age. He continues to have very slow weight gain despite very high calorie diet and maintains a thin body habitus despite this caloric intake.

Ortho: vertical tali bilaterally. Followed by orthopedics and has had casting and surgical repair. Also initially had some overlapping fingers and clenched fists, which has now resolved, though fingers are somewhat long in appearance.

Exam: thin body habitus, elongated head, full cheeks, accessory nipple on right, rocker bottom feet, hypotonia

Development: motor delay due to casting and surgeries on feet. Seems on track socially and verbally. Feeding is improving and takes all feeds orally.

Testing: SNP microarray, normal male
WES: *de novo RARA* variant; c.865G>A, p.Gly289Arg (classified as variant of unknown significance)

Proband 2:

This child was born at 35-4/7 weeks' gestation to a 27-year-old G2P1-2 via urgent C-section. During the pregnancy, his mother was noted to have pelvic kidneys and she underwent nonstress tests twice per week throughout the third trimester, which were all reassuring. She went into labor at 35 weeks and 4 days. When she arrived at the hospital, the baby continued to be in breech presentation with oligohydramnios and therefore a C-section was performed. She noted that fetal monitoring was normal during the C-section itself and during her monitoring time after arrival at the hospital. At birth, he did not cry. He had initial Apgars of 1 at 1 minute, 4 at 5 minutes and 8 at 10 minutes. During this time, he was intubated and placed on an oscillator. He weighed 4 pounds 13 ounces at birth. He was noted at birth to have several anomalies, including clubfeet, overlapping toes and a head shape consistent with persistent breech presentation and an occipital shelf. He was noted to have small ears at that time. He had difficulty weaning off of the ventilator and it was ultimately determined that he needed repeat head imaging. Initial head ultrasound at birth was normal; however, repeat imaging demonstrated a perinatal left MCA stroke that was quite extensive. Since that time, he has had a significant volume loss as well as development of seizures, initially spasms followed by tonic seizures, dysphagia, developmental delay and dependence on tracheostomy with ventilator. His ventilator requirement is likely multifactorial, including chronic respiratory failure secondary to upper airway obstruction, pulmonary hypoplasia, history of tracheal stenosis status post dilation, and severe tracheobronchomalacia. At the age of 3 years he was using the ventilator overnight but it was not needed during the day. He has a diagnosis of spastic hemiplegia, and has global delays. His primary method of ambulation at age 3 years was scooting, and was working on learning to walk independently. He could babble and his primary method of communication was signs and pointing. He was very social and emotionally bonded with his family.

His head shape became progressively brachycephalic in infancy, and CT scan demonstrated bilateral coronal craniosynostosis and bilateral intraoccipital synchondroses (Figure 3). He has not needed cranial vault surgery because his MCA stroke resulted in volume loss (Figure 3). Other medical issues include renal dysplasia with cysts and chronic kidney disease, as well as undescended testes and persistent asymmetric overlapping toes. He was also noted to have esotropia. Moderate to severe conductive hearing loss was noted. He also had an abnormal pattern of tooth eruption.

SNP array was completed and was normal. Trio exome sequencing was subsequently done through GeneDx. He was found to have a *de novo* variant, c.865G>A, p.Gly289Arg, in exon 7 of *RARA*. At the time of the clinical testing, this variant was classified as a variant of uncertain significance.