Small 6q16.1 Deletions Encompassing *POU3F2* Cause Susceptibility to Obesity and Variable Developmental Delay with Intellectual Disability

Paul R. Kasher,[1] Katherine E. Schertz,[2] Megan Thomas,[3,4] Adam Jackson,[3] Silvia Annunziata,[5] María J. Ballesta-Martinez,[6] Philippe M. Campeau,[7] Peter E. Clayton,[8] Jennifer L. Eaton,[9] Tiziana Granata,[5] Encarna Guillén-Navarro,[6] Cristina Hernando,[10] Caroline E. Laverriere,[2] Agne Liedén,[11] Olaya Villa-Marcos,[10] Meriel McEntagart,[11] Ann Nordgren,[12] Chiara Pantaleoni,[5] Céline Pebrel-Richard,[13] Catherine Sarret,[14] Francesca L. Sciacca,[5] Ronnie Wright,[15] Bronwyn Kerr,[1,15] Eric Glasgow,[2] and Siddharth Banka[1,15,*]

Genetic studies of intellectual disability and identification of monogenic causes of obesity in humans have made immense contribution toward the understanding of the brain and control of body mass. The leptin > melanocortin > SIM1 pathway is dysregulated in multiple monogenic human obesity syndromes but its downstream targets are still unknown. In ten individuals from six families, with overlapping 6q16.1 deletions, we describe a disorder of variable developmental delay, intellectual disability, and susceptibility to obesity and hyperphagia. The 6q16.1 deletions segregated with the phenotype in multiplex families and were shown to be de novo in four families, and there was dramatic phenotypic overlap among affected individuals who were independently ascertained without bias from clinical features. Analysis of the deletions revealed a ~350 kb critical region on chromosome 6q16.1 that encompasses a gene for proneuronal transcription factor *POU3F2*, which is important for hypothalamic development and function. Using morpholino and mutant zebrafish models, we show that *POU3F2* lies downstream of *SIM1* and controls oxytocin expression in the hypothalamic neuroendocrine preoptic area. We show that this finding is consistent with the expression patterns of *POU3F2* and related genes in the human brain. Our work helps to further delineate the neuro-endocrine control of energy balance/body mass and demonstrates that this molecular pathway is conserved across multiple species.

Intellectual disability has an estimated prevalence of 1.5%–2.0%[1] and is a genetically and phenotypically heterogeneous group of disorders. Studies of genetic causes of intellectual disability have made immense contributions toward our understanding of the human brain. Obesity and related co-morbidities are a major public health concern across the world.[2] Understanding the control mechanisms of body mass is a fundamental question for biology and an important area for research.

Rare copy-number variations (CNVs) are linked with a range of phenotypes and are a particularly well-recognized cause of developmental disorders and intellectual disability.[3,4] Additionally, rare CNVs can provide insights into the single-gene causes of human disorders[5,6] and can provide clues toward the genetic basis and molecular mechanisms of commoner complex conditions and traits[7,8] including obesity.[9–14] Here, we describe a study of small overlapping 6q16.1 deletions in individuals with variable developmental delay, intellectual disability, and susceptibility to obesity and hyperphagia along with extended analyses that define the likely critical gene for the phenotype and its role in neuro-endocrine control of energy balance and body mass.

Ethics approval for the study was obtained from the NHS ethics committee (11/H1003/3) and the University of Manchester. Informed consent was taken from all participants recruited into the study from the Manchester Centre for Genomic Medicine. Other participants provided consent to publish their data to the recruiting clinician.

We identified a family (referred to as family 1) with four members, most of whom were affected with neonatal hypotonia, gross motor delay, speech delay, intellectual disability, behavioral problems, obesity, and hyperphagia with onset from mid-childhood (Figure 1A). Their clinical features are summarized in Table 1. Family history suggested an autosomal-dominant inheritance pattern, and

[1]Manchester Centre for Genomic Medicine, Institute of Human Development, Faculty of Medical and Human Sciences, University of Manchester, Manchester M13 9WL, UK; [2]Department of Oncology, Georgetown Lombardi Comprehensive Cancer Center, Georgetown University, Washington, DC 20057, USA; [3]Child Health Directorate, Blackpool Teaching Hospitals, Blackpool FY3 8NR, UK; [4]Faculty of Health and Medicine, University of Lancaster, Lancaster LA1 4YW, UK; [5]Fondazione I.R.C.C.S. Istituto Neurologico "C. Besta," Milan 20133, Italy; [6]Sección de Genética Médica, Hospital Clínico Universitario Virgen de la Arrixaca, IMIB-Arrixaca, Cátedra de Genética, UCAM, 30120 Murcia, Spain; [7]Department of Paediatrics, University of Montreal, Montréal, QC H3T 1J4, Canada; [8]Centre for Paediatrics & Child Health, Institute of Human Development, University of Manchester, Manchester M13 9WL, UK; [9]Summa Health System, Akron, OH 44304, USA; [10]Quantitative Genomic Medicine Laboratories (qGenomics), 08950 Barcelona, Spain; [11]Medical Genetics, St George's University Hospitals NHS Foundation Trust, London SW17 0QT, UK; [12]Department of Clinical Genetics, Karolinska University Hospital, Stockholm 171 76, Sweden; Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm 171 76, Sweden; [13]Cytogénétique Médicale, CHU-Estaing, 63003 Clermont-Ferrand, France; [14]Génétique Médicale, Hôpital Estaing, CHU Clermont-Ferrand, 63000 Clermont-Ferrand, France; [15]Manchester Centre for Genomic Medicine, St Mary's Hospital, Central Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre (MAHSC), Manchester M13 9WL, UK
*Correspondence: siddharth.banka@manchester.ac.uk
http://dx.doi.org/10.1016/j.ajhg.2015.12.014.

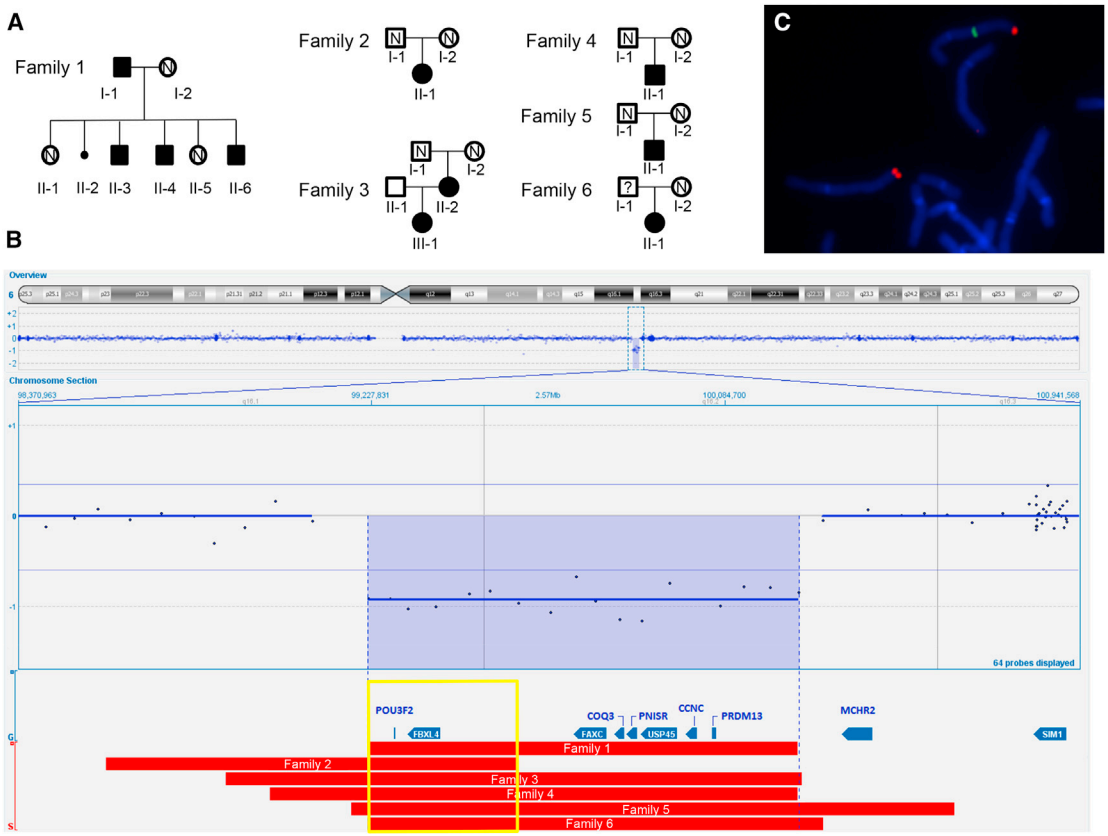


Figure 1. Results of Clinical Genetic Studies

(A) Pedigrees of families. Array comparative genomic hybridization (aCGH) on a DNA sample from individual II-4 of family 1 revealed a 1–1.2 Mb heterozygous deletion on chromosome 6q16.1q16.2 that segregated with the phenotype in the family. We interrogated the local clinical cytogenetics databases of our collaborators for <2 Mb 6q16 deletions that do not include *SIM1* and identified six additional individuals from five families (family 2–6). In four individuals, deletions were proven to have arisen de novo. One individual in family 3 had inherited the deletion from her affected mother. Standard symbols have been used to draw the pedigrees. Dark squares represent affected individuals who were found to have 6q16 deletion. Squares or circles with "N" denote individuals who were tested and found not to have the familial 6q16 deletion. "?" denote individuals whose genotype information is not available.

(B) Results of copy-number analysis. The top panel represents the chromosome bands with the copy-number state of the corresponding hybridized probes from the aCGH results of individual II-4 of family 1. The middle panel focuses on the 6q16 region. The horizontal red bars in the bottom panel show the minimum extent of the microdeletions (in hg19 build) in all five families. The bottom panel is annotated with respective gene loci. The yellow box circumscribes the maximum common overlapping region of the deletion in the five families.

(C) Metaphase fluorescent in situ hybridization (FISH) from individual II-4 from family 1. FISH was undertaken with spectrum green fluorophore-prelabeled RP11-290C18 BAC probe (The Centre for Applied Genomics, Toronto, Canada) which maps to the 6q16.2 region (chr6: 99,813,064–99,990,209). A spectrum orange fluorophore-prelabeled 6q subtelomeric probe (Abbott Molecular) was used as a control. The FISH independently confirmed the heterozygous 6q16.2 deletion in this individual.

Prader Willi syndrome (PW syndrome or PWS) (MIM: 176270) was ruled out via methylation-specific multiplex ligand probe amplification. An array comparative genomic hybridization (aCGH) was performed on a DNA sample from individual II-4 via CytoSure ISCAv2 (8x60k) microarray (Oxford Gene Technology) according to the manufacturer's protocol. CytoSure Interpret v3.4.3 software was used for data analysis, and copy-number aberrations were detected using a minimum of four markers per segment with abnormal log2 ratios (~180 kb backbone and ~15 kb targeted resolution). This revealed a 1–1.2 Mb deletion on chromosome 6q16.1q16.2 (chr6: 99,218,535–100,260,996 in hg19) (DECIPHER: 265018) (Figure 1B). The deletion was absent in the Database of Genomic Variants (DGV) and in more than 6,000 local controls. Metaphase fluorescent in situ hybridization (FISH) analysis performed on lymphocyte cell suspensions using standard protocols independently confirmed the 6q16 deletion in II-4 (Figure 1C) and in all affected family members (II-3, II-6, and I-1). The deletion was absent in all the unaffected siblings. We could not investigate the origin of the deletion in this family because samples from parents of I-1 were not available for testing.

PWS is a classic contiguous gene syndrome that in the majority of affected individuals results from deletion of paternal copies of the imprinted chromosome 15q11–q13 locus.[15] "PW-like syndrome associated with chromosome 6" is another recognized clinical entity (MIM: 176270) that has been described with different genomic deletions of chromosome 6. Of these, some individuals

with interstitial 6q16 deletions most closely resemble the phenotype of PWS. Most such 6q16 deletions encompass *SIM1* (single minded homolog 1 [MIM: 603128]).[9] *SIM1* is part of the central molecular pathway that regulates body mass. In brief, the adipocyte-derived hormone leptin (*LEP* [MIM: 164160]) and its widely expressed leptin receptor (*LEPR* [MIM: 601007]) stimulate proopiomelanocortin (*POMC* [MIM: 176830]) expression in the hypothalamus.[16] POMC is enzymatically cleaved to form α- and β-melanocyte-stimulating hormones, which activate signaling via melanocortin-4 receptor (*MC4R* [MIM: 155541]) to induce expression of *SIM1*. This pathway is dysregulated in multiple monogenic human obesity syndromes (Table S1). Disruptions or heterozygous loss-of-function mutations of *SIM1* cause severe obesity (MIM: 601665).[10–12] *SIM1* is a master regulator of neurogenesis and its optimum dosage is essential for the formation of supraoptic (SON) and paraventricular (PVN) hypothalamic nuclei that play a central role in body mass regulation.[17,18] In mice, *Sim1* haploinsufficiency results in neuroanatomical defects but hyperphagic obesity develops even in the absence of structural abnormalities, which is thought to be mediated by deficiency of oxytocin.[19] However, the mechanism of how *SIM1* regulates oxytocin is not known. There is no known conserved binding site for *Sim1* or its heterodimer partner *Arnt2* in 5 kb upstream or downstream genomic sequence of *Oxt* (oxytocin).[19] Interestingly, some individuals with PWS-like phenotype have 6q16 deletions that do not encompass *SIM1*.[9] This suggests that there is at least one other gene located on 6q16, loss of which can cause intellectual disability and obesity and perhaps lies within the leptin > melanocortin > SIM1 pathway.

The deletion in family 1 encompasses nine known protein-coding genes: *POU3F2* (MIM: 600494), *FBXL4* (MIM: 605654), *FAXC*, *COQ3* (MIM: 605196), *PNISR* (MIM: 616653), *USP45*, *TSTD3*, *CCNC* (MIM: 123838), and *PRDM13* (Figure 1B). Notably, the deletion does not include *SIM1*. Constitutional genomic rearrangements can convey phenotypes through a number of mechanisms including long-range effects.[3] For genomic deletions, the most common mechanism is haploinsufficiency of a single dosage-sensitive critical gene or a group of contiguous genes located within the deleted interval. We, therefore, first investigated the likely effect of haploinsufficiency of the genes within the 6q16.1 deletion. Truncating variants have been described in seven of these nine genes in the general population (Table S2). Loss of one copy of any of these single genes, therefore, is less likely to be driving the congenital or childhood-onset phenotypes in family 1. Out of the remaining two genes, population variant frequency data was unavailable for *TSTD3*, and *POU3F2* was the only gene with no known truncating mutations and low haploinsufficiency index score.[20] This made *POU3F2* an important candidate for further investigation. Another potential gene of interest was *FAXC* because truncating variants in this gene are extremely rare (Table S2).

We interrogated the local clinical cytogenetic databases of our collaborators for <2 Mb 6q16 deletions that excluded *SIM1* and identified six additional individuals from five families (Figure 1; Table S3). Their clinical features were remarkably similar (Table 1). In four individuals, deletions were proven to be de novo in origin and one individual had inherited the deletion from her similarly affected mother (Figure 1).

Overall, in ten individuals (six males and four females) from six families, we report identification of a disorder of developmental delay and intellectual disability with susceptibility to obesity caused by heterozygous 6q16 deletions that encompass *POU3F2* but do not include *SIM1* (Figure 1; Table 1). Our findings are supported by (1) de novo origin of the deletions in at least one affected member of four families; (2) segregation of the phenotype with the deletion in multiplex families; and (3) phenotypic similarity among affected individuals who were independently ascertained without any bias from clinical features. Birth weights of most individuals were within the normal range. Most individuals presented with neonatal hypotonia, although it was not as severe as what is generally encountered in PWS. Unlike PWS, neonatal feeding difficulty, although encountered, was not a major feature in this group of individuals. Most individuals had mild gross and fine motor delay but one individual had normal motor development and another was severely delayed. Most individuals achieved independent sitting between the ages of 6 and 12 months and independent walking between 14 and 21 months. Most individuals had intellectual disability that ranged from mild to moderate. One individual had severe intellectual disability. The body mass index (BMI) could be calculated for four adults in our cohort and it ranged between 3.62 and 4.59 SDs above the mean, putting them in either severely or very severely obese categories. The BMI of all but one child was on or above the 99th centile (range: +1.51 to +4.27 SDs) in the obese category. The BMI of one 13-year-old girl was on the 91st centile, putting her in the overweight category. The age of onset of obesity ranged from the first year of life to mid-teens. All but one affected individual was reported to have abnormally increased appetite. Interestingly, in all the cases where information was available, the onset of obesity preceded hyperphagia. The excess weight in all individuals was more predominant in the truncal area. Relatives frequently described problems with unpredictable behavior with unprovoked outbursts of aggression, tantrums, impulsivity, mood swings, and emotional lability. Some individuals were described as withdrawn and poor at social interaction. Interestingly, a recent genome-wide association study revealed 6q16.1 as a risk locus for bipolar disorder.[21]

A combined analysis of all the deletions revealed that the maximum critical region for the phenotype included only two genes, *POU3F2* and *FBXL4* (Figure 1B). Loss-of-function and truncating *FBXL4* mutations cause autosomal-recessive encephalomyopathic type mitochondrial DNA

Table 1. Clinical Features of Individuals in This Study

Individual#	Pedigree Identifier[a]	Inheritance	Sex	Age	Birth weight in kg (centile)	Neonatal Hypotonia	Feeding Difficulties in Infancy	Motor Delay	Speech Delay	Learning Disability
1	1-I-1	NA	male	61	NA	yes	NA	yes	yes	moderate
2	1-II-3	paternal	male	21	4.8 (98th)	yes	no	mild	yes	mild
3	1-II-4	paternal	male	19	4.3 (91st)	yes	no	mild	yes	moderate
4	1-II-6	paternal	male	10	3.5 (50th)	no	no	mild	yes	moderate
5	2-II-1	DN	female	13	3.4 (50th)	no	no	no	yes	moderate
6	3-II-2	DN	female	52	3.25 (50th)	yes	no	mild	moderate	mild
7	3-III-1	maternal	female	26	3.1 (25th)	yes	no	mild	mild	mild
8	4-II-1	DN	male	9	3.69 (83rd)	yes	no	mild	mild	mild
9	5-II-1	DN	female	4	3.3 (75th)	yes	yes	severe	severe	severe
10	6-II-1	NA	male	16	3.5 (50th)	yes	yes	moderate	severe	moderate

Abbreviations are as follows: DN, de novo; NA, not available; IQ, intelligence quotient; WISC, Wechsler Intelligence Scale for Children.
[a]Pedigree identifier: Family#-Generation#-individual# (compare with Figure 1).
[b]BMI categories for adults are 25 to 30, overweight; 30 to 35, moderately obese; 35 to 40, severely obese; over 40, very severely obese.
[c]BMI categories for children are 85th to 95th centile, overweight; above 95th centile, obese.

depletion syndrome (MIM: 615471) in which heterozygous carriers are phenotypically unaffected.[22] Thus *POU3F2* remained as the most likely critical gene for the phenotype of 6q16.1 deletions.

Class III POU genes, *POU3F1* (MIM: 602479), *POU3F2*, *POU3F3* (MIM: 602480), and *POU3F4* (MIM: 300039), belong to a family of transcription factors that bind to the octameric DNA sequence 5′-ATGCAAAT-3′. These genes share a highly homologous POU domain and are predominantly expressed in the central nervous system. *Pou3f2* and *Pou3f3* upregulate proneuronal genes[23,24] and are required for production, migration, and positioning of neocortical neurons.[25,26] In vivo, *Pou3f2* shares functional redundancy with *Pou3f3* and simultaneous disruption of both genes is required to disturb normal formation of the neocortex and migration of neurons.[26] However, *Pou3f3* single mutants exhibit an abnormal hippocampus[26] and *Pou3f2* single mutants display abnormal neurosecretory neurons of the hippocampal PVN and SON[27,28] thus demonstrating their essential roles in the development of specific areas of the brain.

We examined data from the Human Brain Transcriptome Project[29] that confirmed that *POU3F2* is expressed throughout fetal and adult life in the human brain (Figure S1). We then examined data in the Allen Human Brain Atlas[30] via the Brain Explorer tool for expression patterning of *POU3F2* and related genes in the adult hypothalamus and hippocampus (Figure S2). This confirmed

Behavioral Issues	Latest Weight in kg (centile)	Latest Height in cm (centile)	Latest BMI (SD)	BMI Category	Obesity Onset (years)	Hyperphagia	Hyperphagia Onset (years)	Other Comments
mood swings, occasional aggressiveness	NA	172 (25^{th})	NA	NA	NA	yes	NA	diabetes mellitus at 55 years, angioplasty at 60 years
mood swings, acute depression, occasional aggressiveness	117.8 (>99^{th})	177.3 (50^{th})	37.5 (3.77)	severely obese[b]	15	yes	19	right hydrocele, undescended testis, and inguinal hernia; left convergent squint
withdrawn, occasional aggressiveness	140.0 (>99^{th})	172 (25^{th})	47.3 (4.59)	very severely obese[b]	10	yes	12	left convergent squint
tantrums, emotional liability	59.5 (> 99^{th})	139 (50^{th}–75^{th})	30.8 (4.27)	obese[c]	1	yes	8	left undescended testis, right hydrocele; right talipes; convergent squint
poor social interaction, occasional aggressiveness	57 (75 –90^{th})	160 (75^{th})	22.3 (1.51)	overweight[c]	NA	no	NA	partial regression of speech after 2 years; at 13 years: WISC-III total IQ = 48, verbal IQ = 45; performance IQ = 62
tantrums, rigid behavior	101 (95^{th})	161 (25^{th})	38.9 (3.87)	severely obese[b]	19	yes	21	clumsiness and poor articulation; lordosis; hyperhidrosis.
rigid behavior, occasional aggressiveness	95 (NA)	160 (NA)	37.1 (3.62)	severely obese[b]	11	yes	14	hypothyroidism; systemic lupus erythematous-like illness
impulsiveness, occasional aggressiveness	66 (99^{th})	152.5 (99^{th})	28.38 (3.57)	obese[c]	4	yes	8	IQ = 72 (7.5 years)
sleep problems	24.6 (>97^{th})	109.5 (98^{th})	20.7 (2.99)	obese[c]	2	yes	NA	gastro-esophageal reflux; very poor expressive speech; thin corpus callosum; kyphosis
occasional aggressiveness poor social interaction	88.9 (99^{th})	184 (91^{st})	26.3 (2.29)	obese[c]	10	yes	NA	auditory memory problems and dyspraxia; left cerebellar arachnoid cyst

that *POU3F2* and *POU3F3* are highly expressed in the human hypothalamus and hippocampus, respectively (Figure S2), thus mirroring the known expression pattern in mice.

We also explored the expression of *POU3F2* through quantitative RT-PCR (qRT-PCR) in transformed lymphoblastoid cell lines (LCLs) from affected individual II-6 and his unaffected sibling II-1 in family 1. Amplification of the *POU3F2* transcript occurred at very late cycle numbers, suggesting that this gene is expressed at very low levels in peripheral lymphocytes. We observed an approximate 50% reduction in *POU3F2* expression in LCLs from the affected individual in comparison to cells derived from the unaffected sibling (Figure S3), although the results were not statistically significant (most likely due to the low expression levels resulting in high variability) even after nine technical repeats and by using a high concentration of cDNA template.

We explored the role of *POU3F2* in hypothalamic development using zebrafish models. We determined that zebrafish have two orthologs of human *POU3F2*, *pou3f2a* and *pou3f2b* (Figures S4 and S5 and Table S4). All procedures were in accordance with NIH guidelines on the care and use of animals and were approved by the Georgetown University Institutional Animal Care and Use Committee, Protocol 11-008. Zebrafish (*Danio rerio*) were raised, maintained, and crossed as described previously.[31] Embryos were raised at 28°C and staging was determined by

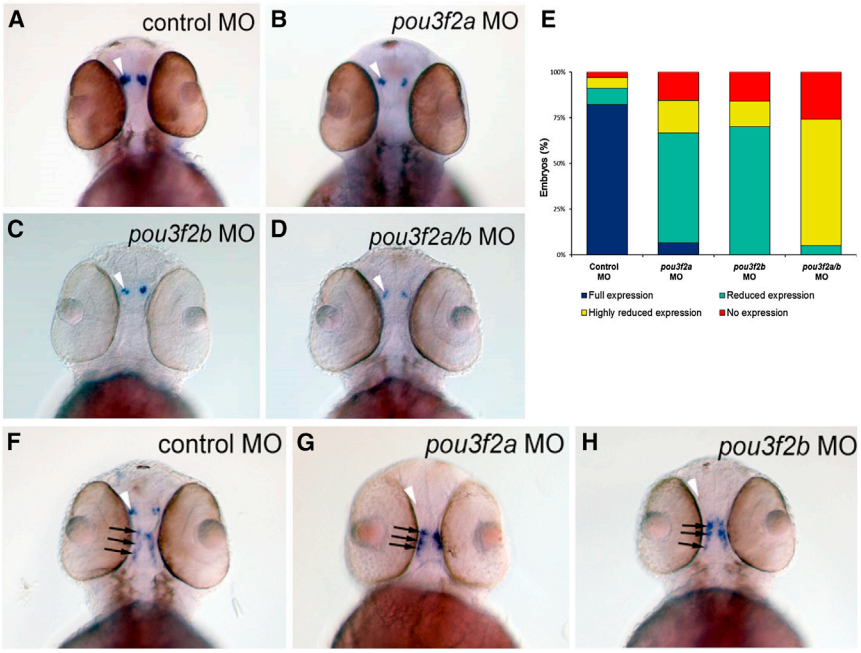


Figure 2. Effect of *pou3f2a* and *pou3f2b* Morpholino Oligonucleotides Knockdown on *oxt*- and *avp*-Expressing Cells

Representative ventral views of 48 hpf embryos stained for *oxt* (A–D) and *avp* (F–H) expression by whole mount in situ hybridization (WISH). The white arrowheads indicate the location of neuroendocrine preoptic area (NPO) and black arrows indicate *avp* expression in the ventral hypothalamus.

(A) Control MO showing full *oxt* expression (n = 67).

(B) *pou3f2a* MO showing reduced *oxt* expression (n = 45).

(C) *pou3f2b* MO showing reduced *oxt* expression (n = 69).

(D) *pou3f2a/pou3f2b* MO-injected embryos showing highly reduced *oxt* expression (n = 96).

(E) Quantification of *oxt* expression. 82% of control MO-injected embryos had full *oxt* expression (blue). Injection of either *pou3f2a* or *pou3f2b* MO resulted in majority of the embryos with reduced *oxt* expression (green). Simultaneous injection of *pou3f2a* and *pou3f2b* MOs resulted in highly reduced *oxt* expression (yellow) majority of the embryos with 26% showing no expression (red).

(F) Control MO showing full *avp* expression within the NPO and ventral hypothalamus.

(G and H) *pou3f2a* (G) and *pou3f2b* (H) MO showing no *avp* expression within the NPO without any reductions in its expression in the ventral hypothalamus.

both hours post fertilization (hpf) and morphological characteristics.[32]

Plasmids for *pou3f2a* and *pou3f2b* were obtained from RZPD (German Science Centre for Genome Research). The *pou3f2a* DIG-labeled antisense riboprobe was generated using T7 polymerase from RsrII linearized plasmid (DIG-labeling Kit, Roche). For *pou3f2b*, the cDNA was subcloned into pBluescript II (Stratagene) and then linearized with KpnI for riboprobe synthesis using T7. Single- and double-labeled whole mount in situ hybridization (WISH) was performed according to previously published protocol.[33] We determined that, in zebrafish, by 48 hpf, *pou3f2a* and *pou3f2b* are normally expressed in the diencephalon, the midbrain tegmentum, and throughout the hindbrain (Figure S6). However, within the diencephalon, strong expression of *pou3f2a* and *pou3f2b* remains restricted to a small area of the neuroendocrine preoptic area (NPO). Double-labeled WISH demonstrated that in the NPO, *pou3f2a* and *pou3f2b* mRNAs normally co-localize in subsets of *oxt*-expressing cells (Figure S7).

Antisense morpholino oligonucleotides (MOs) targeting *pou3f2a* or *pou3f2b* (Table S5) were injected independently or simultaneously into zebrafish embryos at the 1- to 2-cell stage in 1× Danieau's solution at 1.0 ng/embryo. The total amount of injected MO in each group was kept constant for each embryo. *Oxt* and *avp* probes were generated as previously described.[33–35] *Oxt* expression in approximately 30 cells was quantified as full expression, in 5–15 cells as reduced expression, in 1–4 cells as highly reduced expression, and in 0 cells as no expression. Antisense MO-mediated knockdown of *pou3f2a* or *pou3f2b* individually resulted in significantly decreased *oxt* expression (Figure 2). Simultaneous knockdown of *pou3f2a* and *pou3f2b* decreased *oxt* expression further (Figure 2), demonstrating the role of *POU3F2* in regulating *OXT* expression. Likewise, *pou3f2a* or *pou3f2b* MOs individually eliminated *avp* expression within the NPO. However, neither *pou3f2a* nor *pou3f2b* MOs reduced *avp* expression in the ventral hypothalamus, demonstrating the specificity of the MOs.

We have previously shown that MO knockdown of *sim1a* eliminates *oxt* and *avp* expression in zebrafish NPO.[33,34] Additionally, oxytocin expression is decreased in *Sim1* haploinsufficient mice[19] and there is evidence that *Sim1* might regulate *Pou3f2* expression in mice.[17] MO injections for *sim1* were performed as previously published[33,36] and WISH staining for pou3f2a/b expression in the NPO was quantified using an ordinal scale from 0 to 2 as follows: 0, no staining; 1, uncertain or dramatically reduced staining; 2, obvious (normal) staining. This showed significant reduction of *pou3f2a* and *pou3f2b* expression levels in *sim1* morphants (Figures 3A–3D and 3I). *ARNT2* encodes a dimerization partner of *SIM1* for the development of the hypothalamus.[37] In the NPO of previously described[38] homozygous *arnt2*^{hi2639Tg} null zebrafish mutant embryos (Table S6), the expression of *pou3f2a* and *pou3f2b* was undetectable (Figures 3E–3H and 3J), demonstrating the role of *SIM1-ARNT2* dimers in regulating *POU3F2* expression in the hypothalamus. We also examined the expression of *sim1a* in NPO of *pou3f2a* and *pou3f2b* morphants and found no obvious differences from controls (Figure S7).

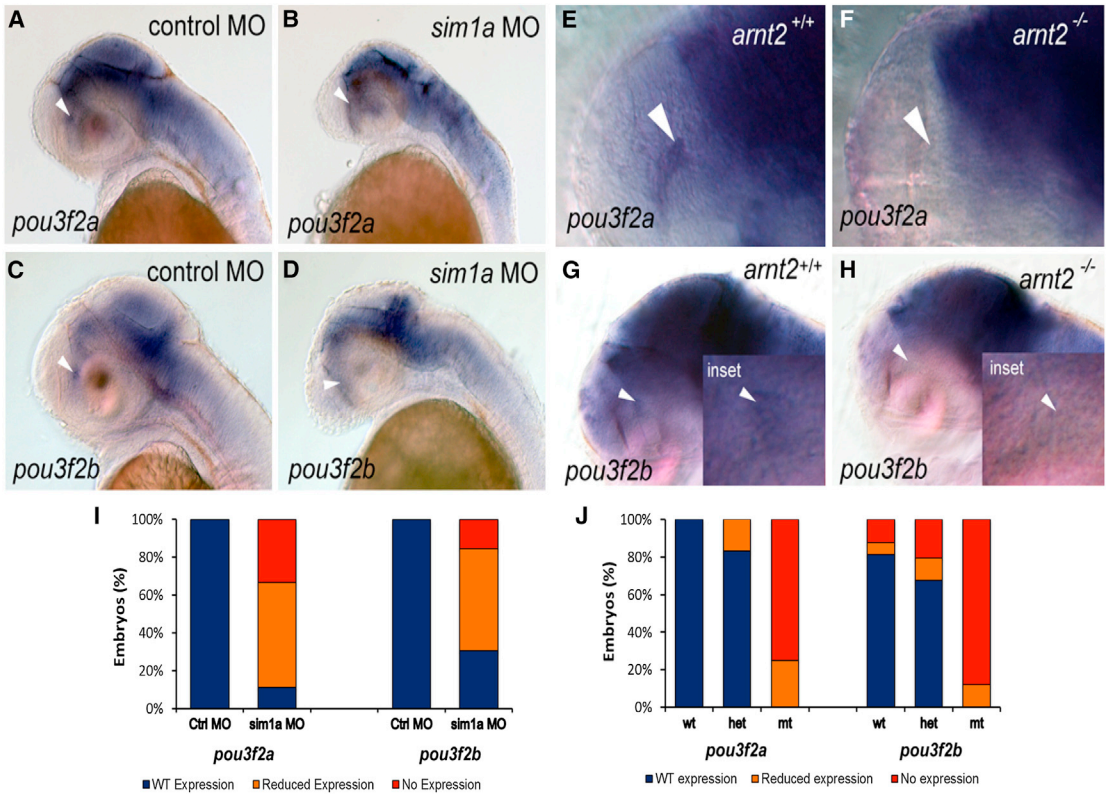


Figure 3. *pou3f2a* and *pou3f2b* Expression Is Reduced in *sim1a* Morphants and Is Eliminated in *arnt2*-Null Mutant Embryos
Representative lateral views of embryos stained for *pou3f2a* or *pou3f2b* expression by WISH at 48 hpf. Eyes have been removed to better visualize the staining in the NPO (indicated by white arrowheads and magnified views are shown in the insets).
(A and C) Control morpholino oligonucleotide (MO)-injected embryos showing normal expression of *pou3f2a* (n = 11) (A) and *pou3f2b* (n = 10) (C).
(B and D) *sim1a* MO knockdown reduces the level of *pou3f2a* (n = 9) (B) and *pou3f2b* (n = 12) (D) expression.
(E and G) Wild-type embryos showing strong *pou3f2a* (n = 7) (E) and *pou3f2b* (n = 13) (G) staining in the NPO.
(F and H) *arnt2*-null mutant embryos showing an absence of *pou3f2a* ($n_{wt} = 7$; $n_{het} = 12$; $n_{hom} = 8$) (F) and *pou3f2b* ($n_{wt} = 13$; n_{het} n = 28; $n_{hom} = 18$) (H) in the NPO.
(I and J) Both *sim1* MO-injected embryos (I) and *arnt2*-null mutants (J) resulted in a significant number of embryos showing reduced (orange) or no (red) expression of *pou3f2a* and *pou3f2b* indicating that their expression in the NPO is dependent on functional sim1a-arnt2 heterodimers.

Overall, the zebrafish experiments showed that *POU3F2* is a downstream target for the SIM1-ARNT2 dimer in the leptin > melanocortin > SIM1 pathway and that *POU3F2* plays an important role in regulating expression of *OXT* in the hypothalamus.

The maximum critical region for the 6q16.1 deletions and our zebrafish work suggests that the phenotype of the individuals described here is due to haploinsufficiency of *POU3F2*. Mice that are homozygous for loss-of-function *Pou3f2* mutations die within 10 days of birth, whereas heterozygous mice have half-the-normal levels of vasopressin and oxytocin in the hypothalamus in comparison with the wild-type mice.[27] This suggests that haploinsufficiency of *POU3F2* might affect hypothalamic development or functions more specifically than other brain regions. The hypothalamus is a critical integrator of neural and humoral signals that has, among its numerous functions, a fundamental role in controlling the body's energy expenditure, food intake, social behavior, learning, and memory. Abnormal development, survival, or function of hypothalamic neurons is known to underpin a number of disorders associated with obesity, hyperphagia, and abnormal neurodevelopment.[16] This further supports the putative role of *POU3F2* deletion in the phenotype seen in the individuals described here. The similarity of the clinical features described here with those caused by loss-of-function *SIM1* mutations[11,12] is in keeping with our conclusions. However, an interesting difference in individuals presented here is that hyperphagia was reported to develop after the onset of obesity. The underlying reason or mechanism for this is not clear and will need identification of younger pre-symptomatic individuals in the future to confirm this observation.

The intellectual disability and neuropsychological problems associated with haploinsufficiency of *POU3F2* and *SIM1* might result from decreased oxytocin levels. Oxytocin is required for activity-dependent cortical development and cortical plasticity[39] and is known to shape social learning[40]

and emotional responses. Future studies of disorders of this pathway could help delineate the precise function of oxytocin in human learning and behavior. Alternatively, the phenotype might also be due to effects that are independent from the role of oxytocin. For example, *POU3F2* is known to regulate *FOXP2*, mutations in which cause speech-language disorder-1 (MIM: 602081).[41,42]

The phenotypic variability among individuals presented here is notable. For example, the phenotype observed in individual 5 (family 2, individual II-1 in Figure 1) is significantly milder than the clinical features of individual 9 (family 5, individual II-1 in Figure 1). The reason behind this is not clear. It could be due to the differences in the size of their deletions—individual 5 has the smallest deletion that includes only two genes (*POU3F2* and *FBXL4*) and individual 9 has the largest deletion that includes *MCHR2* in addition to all the nine genes deleted in family 1. Alternatively, genetic background or environmental factors might influence the phenotype.

It is remarkable that *SIM1* regulates *POU3F2* and, in humans, both genes are located on 6q16.1, less than 1.6 Mb apart. This raised the possibility that *SIM1* and *POU3F2* might share common regulatory regions, deletion of which might have an effect on the 6q16 deletion phenotype. To test this hypothesis, we examined the evolutionary syntenic architecture of the region using Genomicus browser,[43] which showed that *SIM1* and *POU3F2* are located on different chromosomes in a number of species, including mice (Figure S9). This reduces the probability of shared regulatory regions between *SIM1* and *POU3F2* on 6q16.1.

Notably, we have not described any individuals with *POU3F2* point mutations. Hence, a more complex mechanism, such as one involving long-range gene dysregulation, cannot be completely ruled out to underlie the phenotype seen with 6q16.1 deletions. It will be interesting to see whether nonsense or loss-of-function *POU3F2* mutations can result in the same phenotype. This question might be best answered by targeted sequencing or interrogating exome sequencing data from large cohorts of individuals with ID or obesity. We interrogated data from more than 4,000 trios in the DDD study[44] and did not find any truncating *POU3F2* mutation in this cohort. Intellectual disability and obesity both are listed phenotypes in approximately 3% of probands in the DDD study (C. Wright, personal communication). Of note, *POU3F2* is composed of a single exon (Figure S10). Thus, truncating mutations in this gene might not lead to nonsense-mediated decay and, therefore, might not necessarily result in haploinsufficiency. Second, the *POU3F2* sequence is extremely GC rich (Figure S10) and, therefore, accurate sequencing might be challenging. Examination of coverage metrics for *POU3F2* in the ExAC database revealed extremely poor coverage for almost 50% of the gene (Figure S10).

In summary, we have described overlapping 6q16.1 deletions in ten individuals, from six families, with variable developmental delay, intellectual disability, and susceptibility to obesity and hyperphagia. The likely mechanism is haploinsufficiency of *POU3F2*. Our work helps to further define the neuro-endocrine control of energy balance/food intake and its role in human monogenic obesity by demonstrating that *POU3F2* functions downstream to *SIM1* in the leptin > melanocortin > SIM1 > oxytocin pathway and is an important mediator of the clinical and biochemical effects (decreased oxytocin levels) of loss of SIM1 activity. Our zebrafish work, previous work on mouse models,[17–19,27,28] the human phenotypes, and our analysis of the expression patterning of these genes demonstrate that the molecular pathway linking genes related to hypothalamic function is conserved across species, emphasizing its biological importance.

Supplemental Data

Supplemental Data include ten figures and six tables and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2015.12.014.

Acknowledgments

We acknowledge the support of Manchester Biomedical Research Centre, Frimurare Barnhuset i Stockholm, The Swedish Brain Foundation (Hjärnfonden), and the Karolinska Institutet research funds. Zebrafish work was supported in part by NIH/NCI grant CA51008. P.M.C. is funded in part by the Canadian Institutes of Health Research (RN315908 and RN324373) and the Fonds de recherche du Québec en Santé (FRQS 30647). We thank Kit Albaugh for technical assistance with the zebrafish experiments. We thank Karen Marks for assistance with microarray analysis. We thank Yanick J. Crow for providing consumables and equipment for some genetic and zebrafish experiments that are not presented here. We thank Caroline Wright and Rosemary Kelsell from the Sanger Institute for their help in providing the data from the DDD study. We thank the families for taking part in the study and agreeing to publication.

Received: October 6, 2015
Accepted: December 15, 2015
Published: January 28, 2016

Web Resources

The URLs for data presented herein are as follows:

Allen Human Brain Atlas, http://human.brain-map.org/
COBALT: Multiple Alignment Tool, http://www.st-va.ncbi.nlm.nih.gov/tools/cobalt/re_cobalt.cgi
DECIPHER, http://decipher.sanger.ac.uk/
ExAC Browser, http://exac.broadinstitute.org/
Genomicus v80.01, http://www.genomicus.biologie.ens.fr/genomicus-80.01/cgi-bin/search.pl
HBT – Human Brain Transcriptome, http://hbatlas.org/
MUSCLE, http://www.ebi.ac.uk/Tools/msa/muscle/
OMIM, http://www.omim.org/
Synteny Database, http://syntenydb.uoregon.edu/synteny_db/

References

1. Leonard, H., and Wen, X. (2002). The epidemiology of mental retardation: challenges and opportunities in the new millennium. Ment. Retard. Dev. Disabil. Res. Rev. *8*, 117–134.
2. Friedman, J.M. (2000). Obesity in the new millennium. Nature *404*, 632–634.
3. Lupski, J.R., and Stankiewicz, P. (2005). Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. PLoS Genet. *1*, e49.
4. Cooper, G.M., Coe, B.P., Girirajan, S., Rosenfeld, J.A., Vu, T.H., Baker, C., Williams, C., Stalker, H., Hamid, R., Hannig, V., et al. (2011). A copy number variation morbidity map of developmental delay. Nat. Genet. *43*, 838–846.
5. Coe, B.P., Witherspoon, K., Rosenfeld, J.A., van Bon, B.W.M., Vulto-van Silfhout, A.T., Bosco, P., Friend, K.L., Baker, C., Buono, S., Vissers, L.E.L.M., et al. (2014). Refining analyses of copy number variation identifies specific genes associated with developmental delay. Nat. Genet. *46*, 1063–1071.
6. Yagi, H., Furutani, Y., Hamada, H., Sasaki, T., Asakawa, S., Minoshima, S., Ichida, F., Joo, K., Kimura, M., Imamura, S., et al. (2003). Role of TBX1 in human del22q11.2 syndrome. Lancet *362*, 1366–1373.
7. Banka, S., Cain, S.A., Carim, S., Daly, S.B., Urquhart, J.E., Erdem, G., Harris, J., Bottomley, M., Donnai, D., Kerr, B., et al. (2015). Leri's pleonosteosis, a congenital rheumatic disease, results from microduplication at 8q22.1 encompassing GDF6 and SDC2 and provides insight into systemic sclerosis pathogenesis. Ann. Rheum. Dis. *74*, 1249–1256.
8. Antonarakis, S.E., and Beckmann, J.S. (2006). Mendelian disorders deserve more attention. Nat. Rev. Genet. *7*, 277–282.
9. Rosenfeld, J.A., Amrom, D., Andermann, E., Andermann, F., Veilleux, M., Curry, C., Fisher, J., Deputy, S., Aylsworth, A.S., Powell, C.M., et al. (2012). Genotype-phenotype correlation in interstitial 6q deletions: a report of 12 new cases. Neurogenetics *13*, 31–47.
10. Holder, J.L., Jr., Butte, N.F., and Zinn, A.R. (2000). Profound obesity associated with a balanced translocation that disrupts the SIM1 gene. Hum. Mol. Genet. *9*, 101–108.
11. Bonnefond, A., Raimondo, A., Stutzmann, F., Ghoussaini, M., Ramachandrappa, S., Bersten, D.C., Durand, E., Vatin, V., Balkau, B., Lantieri, O., et al. (2013). Loss-of-function mutations in SIM1 contribute to obesity and Prader-Willi-like features. J. Clin. Invest. *123*, 3037–3041.
12. Ramachandrappa, S., Raimondo, A., Cali, A.M.G., Keogh, J.M., Henning, E., Saeed, S., Thompson, A., Garg, S., Bochukova, E.G., Brage, S., et al. (2013). Rare variants in single-minded 1 (SIM1) are associated with severe obesity. J. Clin. Invest. *123*, 3042–3050.
13. Han, J.C., Liu, Q.-R., Jones, M., Levinn, R.L., Menzie, C.M., Jefferson-George, K.S., Adler-Wailes, D.C., Sanford, E.L., Lacbawan, F.L., Uhl, G.R., et al. (2008). Brain-derived neurotrophic factor and obesity in the WAGR syndrome. N. Engl. J. Med. *359*, 918–927.
14. Mou, Z., Hyde, T.M., Lipska, B.K., Martinowich, K., Wei, P., Ong, C.-J., Hunter, L.A., Palaguachi, G.I., Morgun, E., Teng, R., et al. (2015). Human Obesity Associated with an Intronic SNP in the Brain-Derived Neurotrophic Factor Locus. Cell Rep. *13*, 1073–1080.
15. Ledbetter, D.H., Riccardi, V.M., Airhart, S.D., Strobel, R.J., Keenan, B.S., and Crawford, J.D. (1981). Deletions of chromosome 15 as a cause of the Prader-Willi syndrome. N. Engl. J. Med. *304*, 325–329.
16. Ramachandrappa, S., and Farooqi, I.S. (2011). Genetic approaches to understanding human obesity. J. Clin. Invest. *121*, 2080–2086.
17. Michaud, J.L., Rosenquist, T., May, N.R., and Fan, C.-M. (1998). Development of neuroendocrine lineages requires the bHLH-PAS transcription factor SIM1. Genes Dev. *12*, 3264–3275.
18. Duplan, S.M., Boucher, F., Alexandrov, L., and Michaud, J.L. (2009). Impact of Sim1 gene dosage on the development of the paraventricular and supraoptic nuclei of the hypothalamus. Eur. J. Neurosci. *30*, 2239–2249.
19. Kublaoui, B.M., Gemelli, T., Tolson, K.P., Wang, Y., and Zinn, A.R. (2008). Oxytocin deficiency mediates hyperphagic obesity of Sim1 haploinsufficient mice. Mol. Endocrinol. *22*, 1723–1734.
20. Huang, N., Lee, I., Marcotte, E.M., and Hurles, M.E. (2010). Characterising and predicting haploinsufficiency in the human genome. PLoS Genet. *6*, e1001154.
21. Mühleisen, T.W., Leber, M., Schulze, T.G., Strohmaier, J., Degenhardt, F., Treutlein, J., Mattheisen, M., Forstner, A.J., Schumacher, J., Breuer, R., et al. (2014). Genome-wide association study reveals two new risk loci for bipolar disorder. Nat. Commun. *5*, 3339.
22. Gai, X., Ghezzi, D., Johnson, M.A., Biagosch, C.A., Shamseldin, H.E., Haack, T.B., Reyes, A., Tsukikawa, M., Sheldon, C.A., Srinivasan, S., et al. (2013). Mutations in FBXL4, encoding a mitochondrial protein, cause early-onset mitochondrial encephalomyopathy. Am. J. Hum. Genet. *93*, 482–495.
23. Castro, D.S., Skowronska-Krawczyk, D., Armant, O., Donaldson, I.J., Parras, C., Hunt, C., Critchley, J.A., Nguyen, L., Gossler, A., Göttgens, B., et al. (2006). Proneural bHLH and Brn proteins coregulate a neurogenic program through cooperative binding to a conserved DNA motif. Dev. Cell *11*, 831–844.
24. Dominguez, M.H., Ayoub, A.E., and Rakic, P. (2013). POU-III Transcription Factors (Brn1, Brn2, and Oct6) Influence Neurogenesis, Molecular Identity, and Migratory Destination of Upper-Layer Cells of the Cerebral Cortex. Cereb. Cortex *23*, 2632–2643.
25. McEvilly, R.J., de Diaz, M.O., Schonemann, M.D., Hooshmand, F., and Rosenfeld, M.G. (2002). Transcriptional regulation of cortical neuron migration by POU domain factors. Science *295*, 1528–1532.
26. Sugitani, Y., Nakai, S., Minowa, O., Nishi, M., Jishage, K., Kawano, H., Mori, K., Ogawa, M., and Noda, T. (2002). Brn-1 and Brn-2 share crucial roles in the production and positioning of mouse neocortical neurons. Genes Dev. *16*, 1760–1765.
27. Nakai, S., Kawano, H., Yudate, T., Nishi, M., Kuno, J., Nagata, A., Jishage, K., Hamada, H., Fujii, H., Kawamura, K., et al. (1995). The POU domain transcription factor Brn-2 is required for the determination of specific neuronal lineages in the hypothalamus of the mouse. Genes Dev. *9*, 3109–3121.
28. Schonemann, M.D., Ryan, A.K., McEvilly, R.J., O'Connell, S.M., Arias, C.A., Kalla, K.A., Li, P., Sawchenko, P.E., and Rosenfeld, M.G. (1995). Development and survival of the endocrine hypothalamus and posterior pituitary gland requires the neuronal POU domain factor Brn-2. Genes Dev. *9*, 3122–3135.
29. Kang, H.J., Kawasawa, Y.I., Cheng, F., Zhu, Y., Xu, X., Li, M., Sousa, A.M.M., Pletikos, M., Meyer, K.A., Sedmak, G., et al.

(2011). Spatio-temporal transcriptome of the human brain. Nature *478*, 483–489.
30. Hawrylycz, M., Ng, L., Feng, D., Sunkin, S., Szafer, A., and Dang, C. (2014). The Allen Brain Atlas. In Springer Handbook of Biol, N.K. Neuroinformatics, ed. (Springer Berlin Heidelberg), pp. 1111–1126.
31. Westerfield, M. (1995). The Zebrafish Book (Eugene, Oregon: University of Oregon Press).
32. Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B., and Schilling, T.F. (1995). Stages of embryonic development of the zebrafish. Dev. Dyn. *203*, 253–310.
33. Eaton, J.L., and Glasgow, E. (2006). The zebrafish bHLH PAS transcriptional regulator, single-minded 1 (sim1), is required for isotocin cell development. Dev. Dyn. *235*, 2071–2082.
34. Eaton, J.L., Holmqvist, B., and Glasgow, E. (2008). Ontogeny of vasotocin-expressing cells in zebrafish: selective requirement for the transcriptional regulators orthopedia and single-minded 1 in the preoptic area. Dev. Dyn. *237*, 995–1005.
35. Unger, J.L., and Glasgow, E. (2003). Expression of isotocin-neurophysin mRNA in developing zebrafish. Gene Expr. Patterns *3*, 105–108.
36. Eaton, J.L., and Glasgow, E. (2007). Zebrafish orthopedia (otp) is required for isotocin cell development. Dev. Genes Evol. *217*, 149–158.
37. Michaud, J.L., DeRossi, C., May, N.R., Holdener, B.C., and Fan, C.-M. (2000). ARNT2 acts as the dimerization partner of SIM1 for the development of the hypothalamus. Mech. Dev. *90*, 253–261.
38. Golling, G., Amsterdam, A., Sun, Z., Antonelli, M., Maldonado, E., Chen, W., Burgess, S., Haldi, M., Artzt, K., Farrington, S., et al. (2002). Insertional mutagenesis in zebrafish rapidly identifies genes essential for early vertebrate development. Nat. Genet. *31*, 135–140.
39. Zheng, J.-J., Li, S.-J., Zhang, X.-D., Miao, W.-Y., Zhang, D., Yao, H., and Yu, X. (2014). Oxytocin mediates early experience-dependent cross-modal plasticity in the sensory cortices. Nat. Neurosci. *17*, 391–399.
40. Carter, C.S. (2014). Oxytocin pathways and the evolution of human behavior. Annu. Rev. Psychol. *65*, 17–39.
41. Lai, C.S.L., Fisher, S.E., Hurst, J.A., Vargha-Khadem, F., and Monaco, A.P. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. Nature *413*, 519–523.
42. Maricic, T., Günther, V., Georgiev, O., Gehre, S., Ćurlin, M., Schreiweis, C., Naumann, R., Burbano, H.A., Meyer, M., Lalueza-Fox, C., et al. (2013). A Recent Evolutionary Change Affects a Regulatory Element in the Human FOXP2 Gene. Mol. Biol. Evol. *30*, 844–852.
43. Louis, A., Nguyen, N.T.T., Muffato, M., and Roest Crollius, H. (2015). Genomicus update 2015: KaryoView and MatrixView provide a genome-wide perspective to multispecies comparative genomics. Nucleic Acids Res. *43*, D682–D689.
44. Fitzgerald, T.W., Gerety, S.S., Jones, W.D., van Kogelenberg, M., King, D.A., McRae, J., Morley, K.I., Parthiban, V., Al-Turki, S., Ambridge, K., et al.; Deciphering Developmental Disorders Study (2015). Large-scale discovery of novel genetic causes of developmental disorders. Nature *519*, 223–228.

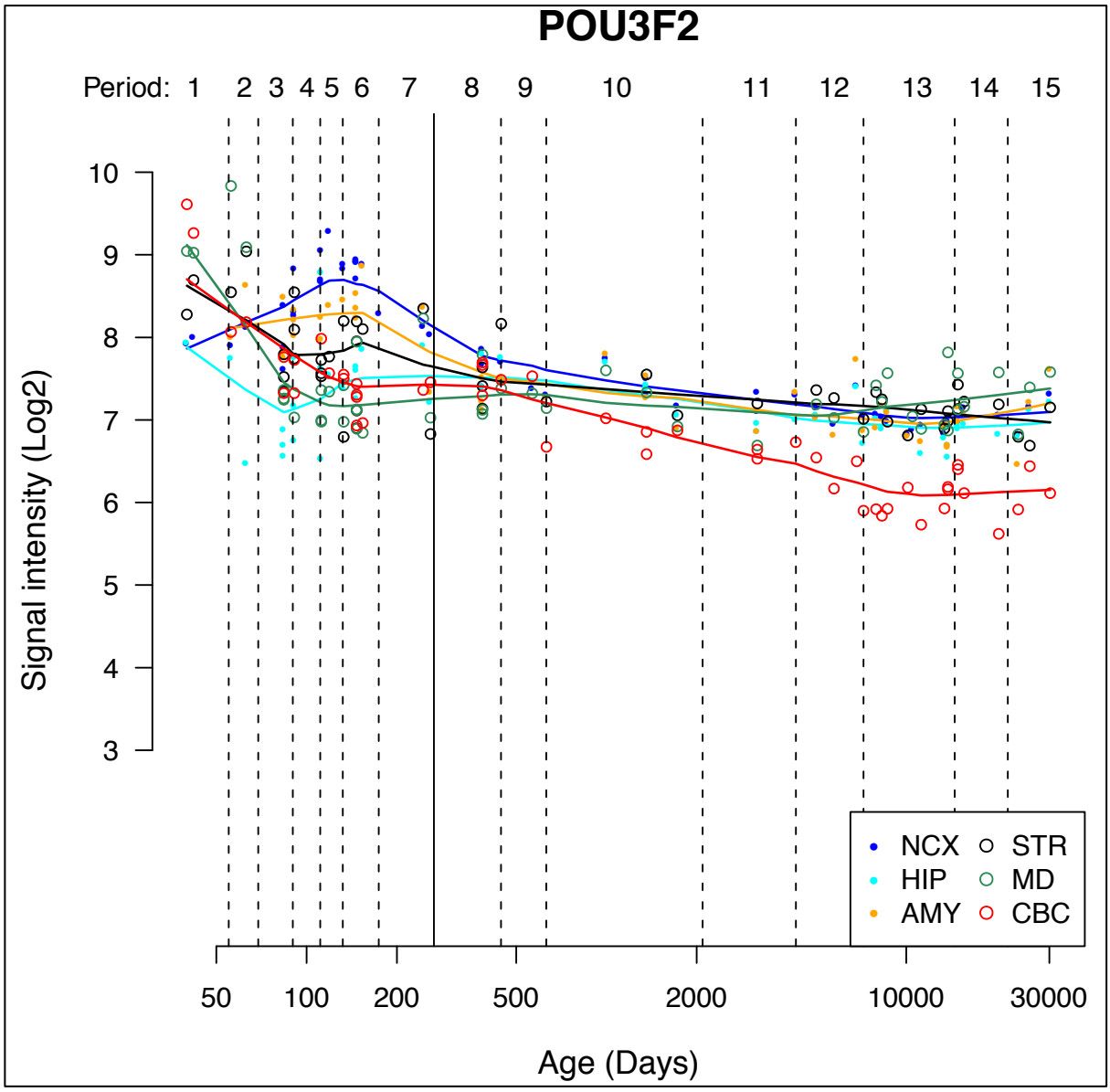
The American Journal of Human Genetics

Supplemental Data

Small 6q16.1 Deletions Encompassing *POU3F2* Cause Susceptibility to Obesity and Variable Developmental Delay with Intellectual Disability

Paul R. Kasher, Katherine E. Schertz, Megan Thomas, Adam Jackson, Silvia Annunziata, María J. Ballesta-Martinez, Philippe M. Campeau, Peter E. Clayton, Jennifer L. Eaton, Tiziana Granata, Encarna Guillén-Navarro, Cristina Hernando, Caroline E. Laverriere, Agne Liedén, Olaya Villa-Marcos, Meriel McEntagart, Ann Nordgren, Chiara Pantaleoni, Céline Pebrel-Richard, Catherine Sarret, Francesca L. Sciacca, Ronnie Wright, Bronwyn Kerr, Eric Glasgow, and Siddharth Banka

Figure S1. *POU3F2* is highly expressed throughout fetal and adult human brain.
Figure generated from Human Brain Transcriptome project data (http://hbatlas.org/) [1] showing high spatio-temporal expression of *POU3F2*.
The legends and axes are labeled and are self-explanatory.



Key: *NCX neocortex; HIP hippocampus; AMY amygdala; STR striatum; MD Mediodorsal nucleus of the thalamus; CBC cerebellar cortex.*

Figure S2: High level of expression of *POU3F2* and related genes in the human hypothalamus.

These figures were generated using Brain Explorer tool based on data from the Allen Human Brain Atlas (http://human.brain-map.org/) [2] in following steps –

(i) Brain Explorer tool was downloaded from http://human.brain-map.org/static/brainexplorer and installed on a local computer.

(ii) Allen Human Brain Atlas (37MB) – October 2013 was installed from http://www.brain-map.org

(iii) Gene name was entered in the search box at URL - http://human.brain-map.org/microarray/search (with settings: Resolution = structures; Color Map = z-score)

(iv) Relevant heat map was selected by clicking that activates the Brain Explorer link on the webpage.

(v) Clicking on the activated link Brain Explorer link opens the Application and enables visualization of spatial rendition of gene expression.

(vi) Within the application specific brain structures of interest were selected to be highlighted and to restrict the visualized expression points (in this case the hypothalamus and the hippocampal formation).

(vii) Various brain sections were visualized and a representative image was selected and exported as an image file.

Figures are arranged to represent the Leptin>MC4R>SIM1>POU3F2>OXT pathway and show that all five genes are highly expressed in the human hypothalamus. The bottom left figure, represents expression pattern of *POU3F3*. Comparing expression patterns of human *POU3F3* and *POU3F2* shows that it is similar to what has been observed in mice [3].

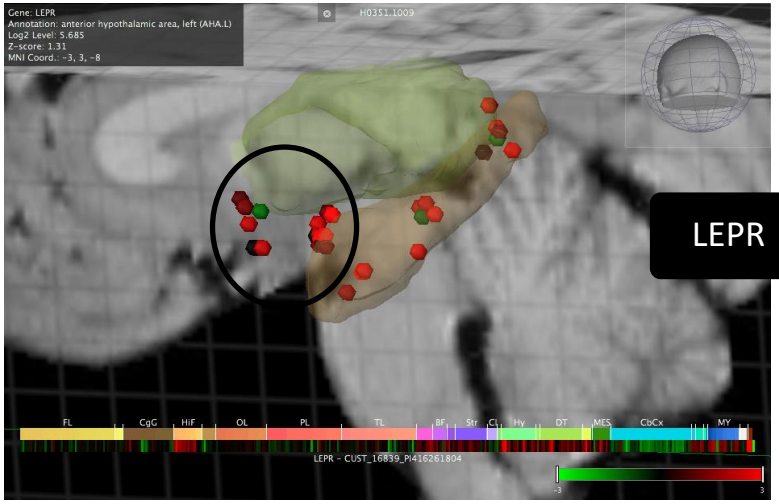
In all the figures the inset at the top left corner provides the gene name, anatomical centre of the figure, Log2 level, Z-score and co-ordinates.

The top right corner inset shows a 'compass' to aid orientation of the main figure.

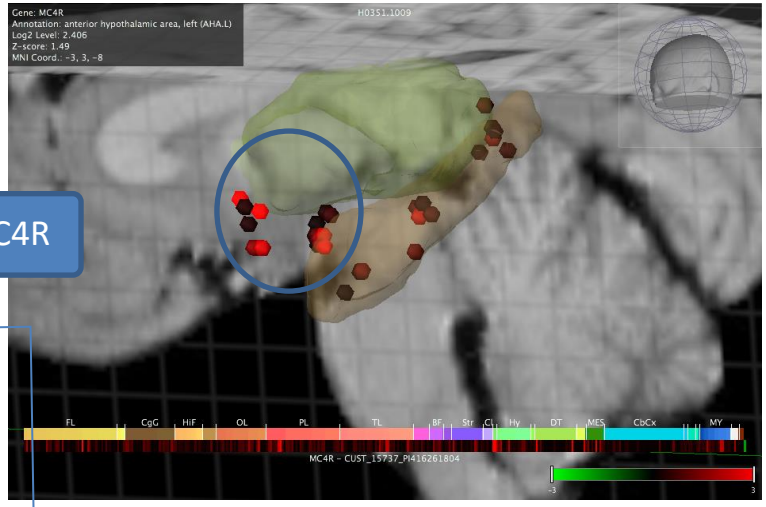
The hypothalamic area is circled, transparent green area represents the thalamus and transparent orange area represents the hippocampal formation.

The expression level at each measured point is given on a green-red scale (bottom right). The two horizontal bars at the bottom provide more detailed information on the expression of the particular gene via the green-red heat map in each of the following regions of the brain – FL (frontal lobe), CgC (cingulate gyrus), HiF (hippocampal formation), OL (occipital lobe), PL (parietal lobe), TL (temporal lobe), BF (basal forebrain), Str (striatum), C (claustrum), Hy (hypothalamus), DT (dentate nucleus), MES (mesencephalon), CbCx (cerebellar cortex) and My (myelencephalon).

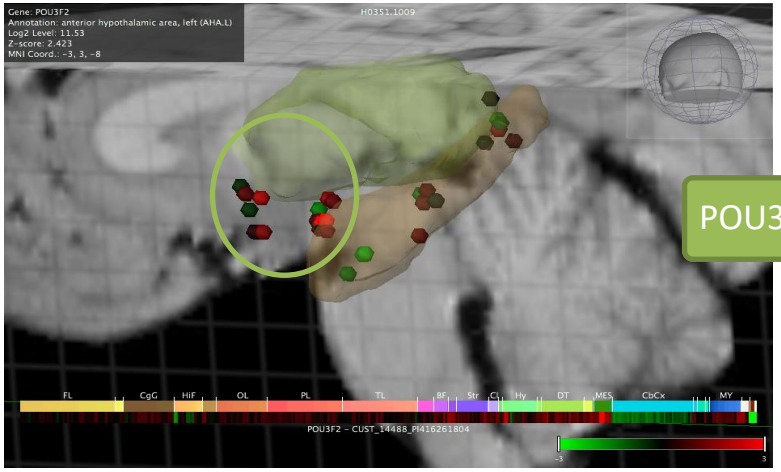
The anatomical section and the donor sample identifiers are provided at the top centre and bottom centre respectively.



LEPR

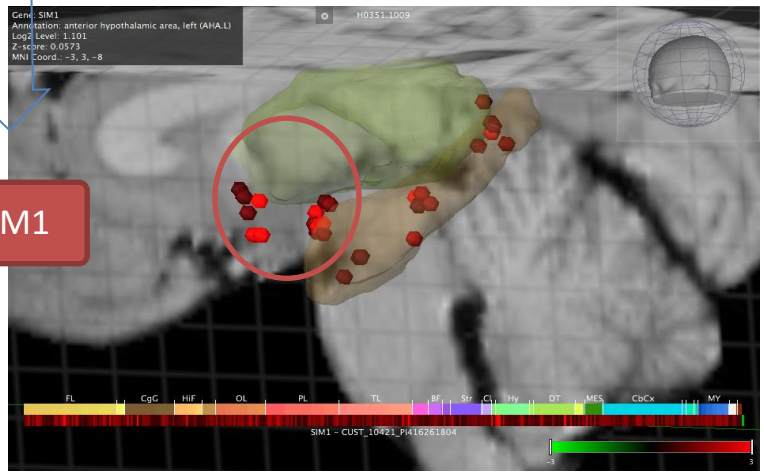


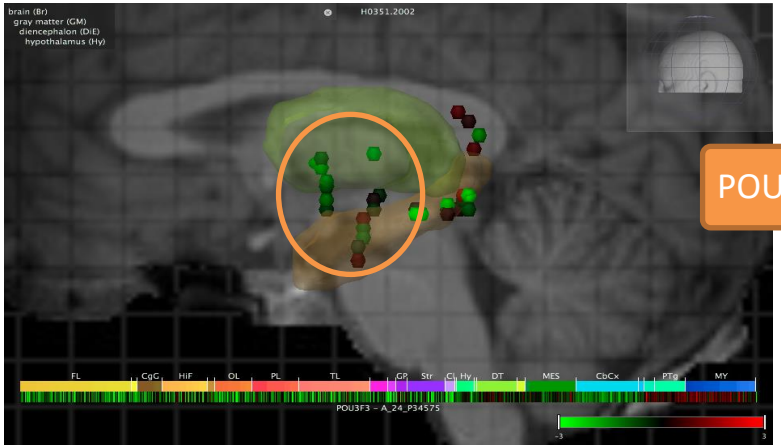
MC4R



POU3F2

SIM1





POU3F3

OXT

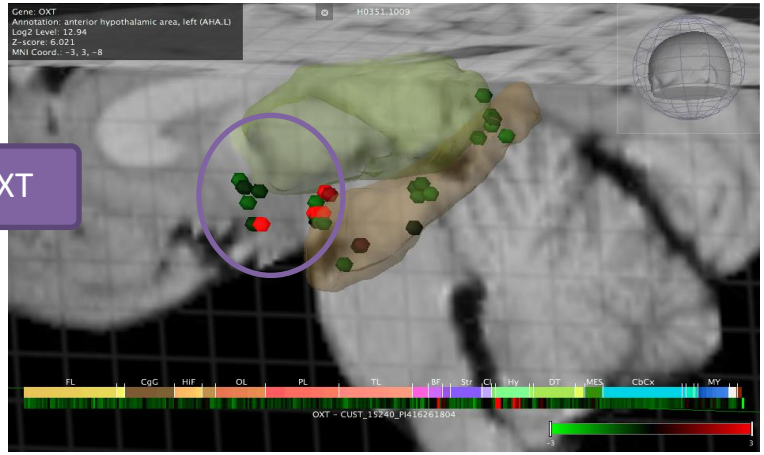
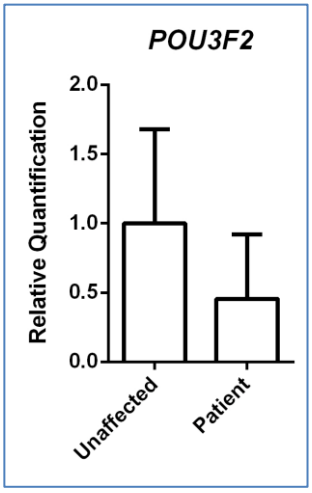


Figure S3: *POU3F2* quantitative reverse transcription PCR in patient cell lines.

POU3F2 expression was assessed in Epstein-Barr virus transformed lymphoblastoid cell lines (LCLs) derived from a patient and unaffected sibling (Family 1). LCLs were maintained as described previously[4]. Total RNA was extracted from LCLs ($10x10^6$ cells per sample) using TRIzol reagent (Thermo Fisher). RNA concentration was assessed using a spectrophotometer (FLUOstar Omega, Labtech). cDNA was synthesised from 1600ng RNA using the High Capacity RNA to cDNA kit (Life Technologies, UK). Quantitative reverse transcription PCR (qRT-PCR) analysis was performed using the TaqMan Universal PCR Master Mix (Applied Biosystems). Each PCR was performed three times in triplicate per cDNA sample (total of 9 technical repeats per sample). The relative abundance of *POU3F2* transcript was measured using a *POU3F2* Taqman probe (Hs00271595_s1) following normalisation to the expression level of two 'housekeeper' genes, *HPRT1* (Hs03929096_g1) and *18s* (Hs999999001_s1), and assessed with the Applied Biosystems StepOne Software v2.1. Statistical significance between groups was determined by t-tests using DataAssist v2.0 (Applied Biosystems). Fold change values were plotted using Graphpad Prism v5.0 and defined as the Relative Quantification (RQ).



An approximate 50% reduction in POU3F2 expression in LCLs derived from patient in comparison to his unaffected sibling (p=0.1447).
Amplification of the POU3F2 transcript occurred at very late cycle numbers, suggesting that this gene is expressed at very low levels in peripheral lymphocytes.
Indeed, we hypothesise that low expression is the most likely cause for the high variation observed, even after 9 technical repeats and using a high concentration of cDNA template.

Figure S4: Human and zebrafish POU3F2 genes are highly conserved.

Protein sequence alignment was created using Colbalt Constraint-based Multiple Protein Alignment Tool. The sequences correspond to zebrafish Pou3f2a (NP_571364), zebrafish Pou3f2b (NP_571235.1), human POU3F2 (accession number: NP_005595), and mouse POU3F2* (accession number: NP_032925.1), human POU3F1 (NP_002690), human POU3F3 (NP_006227), and human POU3F4 (NP_000298).

The entire protein sequence is highly conserved among mice, humans, and both orthologs in zebrafish. The POU homeodomain (amino acids 95-251; Pou3f2a sequence), highlighted in red, has 92% identity and 97% similarity for all seven sequences.

FIGURE 3

```
Pou3f2a1     ----LSH---   -----------   ------GgEGSPWPA   SPLGEQDIKPV-E---E------------LQ   H  28
Pou3f2b1     ---ALSH---   -----------   ------G-EGGPWSS   SPLGEQDIKPA-V---QSPR-DEMHNSSNLQ   H  39
POU3F2 1     witALSHGGG   GGGGGGGGGGG   GGGGGGG-DGSPWST   SPLGQPDIKPS-VvVQQGGRGDELHGPGALQ   Q  66
POU3F2*1     witALSHGGG   GGGGGGGGGGG   GGGGGGG-DGSPWST   SPLGQPDIKPS-VvVQQGGRGDELHGPGALQ   Q  66
POU3F1 1     ---GAGHPVG[9]GGGGGGDWAGG[5]GKAGGGG-TGR----   --------------ADDGGGGGGFHA-----   -  54
POU3F3 1     witALSHGGG   GGGGGGGGGGG   GGGGGGG-DGSPWST   SPLGQPDIKPS-VvVQQGGRGDELHGPGALQ   Q  66
POU3F4 1     ----------   -----------   -----LS-DGGPWSS[4]SPLDQQDVKPGrE---D------------LQ[7]R  37

Pou3f2a29    S    P--RQAHLVPQSQ--HHETAAWRATTTA-HMP-SMSTSNGqSL   IYSQP---   --------GY--G--EMH  77
Pou3f2b40    Q    S--RPPHLVHQTHGNHHDSRAWRTTTAA-HIP-SMATSNGqSL   IYSQPSFS   VNGLIPGSGQ--G--IHH  101
POU3F2 67    Q[21]Q--RPPHLVHHAANHHPGPGAWRSAAAAAHLPPSMGASNG-GL   LYSQPSFT   VNGMLGAGGQPAG--LHH  152
POU3F2*67    Q[21]QqqRPPHLVHHAANHHPGPGAWRSAAAAAHLPPSMGASNG-GL   LYSQPSFT   VNGMLGAGGQPAG--LHH  154
POU3F1 55    -    ------RLVHQGAAH--AGAAWAQGSTAHHLGPAMSPSPG-AS[9]LYAQAAYP[6]LAGMLAAGGGGAGpgLHH  129
POU3F3 67    Q[21]Q--RPPHLVHHAANHHPGPGAWRSAAAAAHLPPSMGASNG-GL   LYSQPSFT   VNGMLGAGGQPAG--LHH  152
POU3F4 38    S    P--HVAHHSPHTN--HPN--AWGASPAP-N-P-SI-TSSGqPL[1]VYSQP---   --------GF--T--VSG  83

Pou3f2a78    H------EEhhS    PHLSE---HG   HPQSLH---------------------------SDEDTPTSDDLEQFAKQF  114
Pou3f2b102   HSMRDAHEDhhS    PHLSD---HG   HPPSQHQHQSHQ-------------SHHD---HSDEDTPTSDDLEQFAKQF  155
POU3F2 153   HGLRDAHDE---    PHHAD---HH   PHPHSHPHQQPPPPPPPQGPPGHPGAHHD--PHSDEDTPTSDDLEQFAKQF  217
POU3F2*155   HGLRDAHDE---    PHHAD---HH   PHPHSHPHQQPPPPPPPQGPPGHPGAHHD--PHSDEDTPTSDDLEQFAKQF  219
POU3F1 130   ALHEDGHEA--Q[ 5]PPHLG---AH   GHAHGHAHAGGLHAAAAHLHPGAGGGGSSvgEHSDEDAPSSDDLEQFAKQF  202
POU3F3 153   HGLRDAHDE---    PHHAD---HH   PHPHSHPHQQPPPPPPPQGPPGHPGAHHD--PHSDEDTPTSDDLEQFAKQF  217
POU3F4 84    M------LEhgG[14]PVLREppdHG[4]HHCQDH---------------------------SDEETPTSDELEQFAKQF  141

Pou3f2a115   KQRRIKLGFTQADVGLALGTLYGNVFSQTTICRFEALQLSFKNMCKLKPLLNKWLEEADSTSGSPTSLDKIAAQGRKRKK  194
Pou3f2b156   KQRRIKLGFTQADVGLALGTLYGNVFSQTTICRFEALQLSFKNMCKLKPLLNKWLEEADSTSGSPTSLDKIAAQGRKRKK  235
POU3F2 218   KQRRIKLGFTQADVGLALGTLYGNVFSQTTICRFEALQLSFKNMCKLKPLLNKWLEEADSSSGSPTSIDKIAAQGRKRKK  297
POU3F2*220   KQRRIKLGFTQADVGLALGTLYGNVFSQTTICRFEALQLSFKNMCKLKPLLNKWLEEADSSSGSPTSIDKIAAQGRKRKK  299
POU3F1 203   KQRRIKLGFTQADVGLALGTLYGNVFSQTTICRFEALQLSFKNMCKLKPLLNKWLEETDSSSGSPTNLDKIAAQGRKRKK  282
POU3F3 218   KQRRIKLGFTQADVGLALGTLYGNVFSQTTICRFEALQLSFKNMCKLKPLLNKWLEEADSSSGSPTSIDKIAAQGRKRKK  297
POU3F4 142   KQRRIKLGFTQADVGLALGTLYGNVFSQTTICRFEALQLSFKNMCKLKPLLNKWLEEADSSTGSPTSIDKIAAQGRKRKK  221

Pou3f2a195   RTSIEVSVKGALESHFLKCPKPGASEINSLADSLQLEKEVVRVWFCNRRQKEKRMTPQNGP-MAGSEDVY   G---DAS  267
Pou3f2b236   RTSIEVSVKGALESHFLKCPKPAASEITSLADSLQLEKEVVRVWFCNRRQKEKRMTPPGGP-LPGTEDVY   G---DTP  308
POU3F2 298   RTSIEVSVKGALESHFLKCPKPSAQEITSLADSLQLEKEVVRVWFCNRRQKEKRMTPPGGT-LPGAEDVY   GGSRDTP  373
POU3F2*300   RTSIEVSVKGALESHFLKCPKPSAQEITSLADSLQLEKEVVRVWFCNRRQKEKRMTPPGGT-LPGAEDVY   GGSRDTP  375
POU3F1 283   RTSIEVGVKGALESHFLKCPKPSAHEITGLADSLQLEKEVVRVWFCNRRQKEKRMTPAAGAgHPPMDDVY[7]GGGGASP  366
POU3F3 298   RTSIEVSVKGALESHFLKCPKPSAQEITSLADSLQLEKEVVRVWFCNRRQKEKRMTPPGGT-LPGAEDVY   GGSRDTP  373
POU3F4 222   RTSIEVSVKGVLETHFLKCPKPAAQEISSLADSLQLEKEVVRVWFCNRRQKEKRMTPP------------   G---DQQ  283

Pou3f2a268   P    H   HGAQTPVP     277
Pou3f2b309   P    H   HGVQTPVQ     318
POU3F2 374   P    H   HGVQTPVQ     383
POU3F2*376   P    H   HGVQTPVQ     385
POU3F1 367   P[11]H[4]HTLPGSVQ     391
POU3F3 374   P    H   HGVQTPVQ     383
POU3F4 284   P    H[4]HTVKTDTS[4]  301

*POU3F2 mouse
```

Figure S5: Human POU3F2 is orthologous with zebrafish Pou3f2a and Pou3f2b.
A phylogeny tree was created using MUSCLE Multiple Sequence Alignment. Protein sequences for all four human POU3 family proteins, POU3F1 (NP_002690), POU3F2 (NP_005595), POU3F3 (NP_006227), POU3F4 (NP_000298), as well as sequences for zebrafish Pou3f2a (accession number: NP_571364) and Pou3f2b (accession number: NP_571235.1), were used as input data. The output data cluster human POU3F2 more closely with zebrafish pou3f2a and pou3f2b than the other POU3 family proteins confirming that these genes are orthologous.

Human *POU3F2*, and zebrafish *pou3f2a,* and *pou3f2b* contain single protein coding exons, as well as conserved synteny (Synteny Database, http://syntenydb.uoregon.edu) (data not shown). The conserved synteny of human *POU3F2* with both *pou3f2a* and *pou3f2b* gene regions indicates that chromosomal duplication likely generated the zebrafish *pou3f2* paralogs.

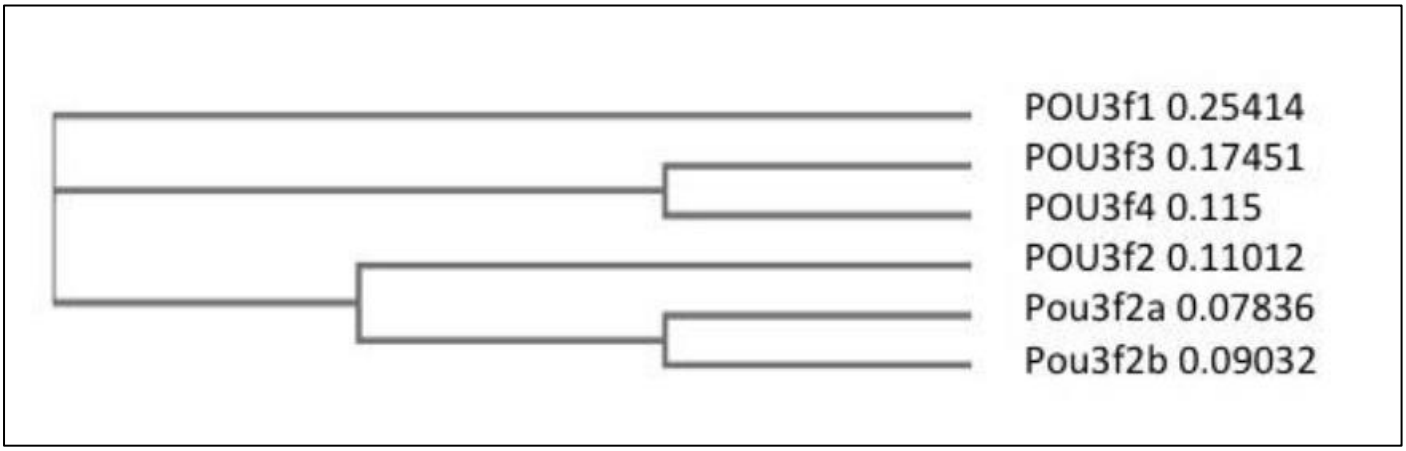
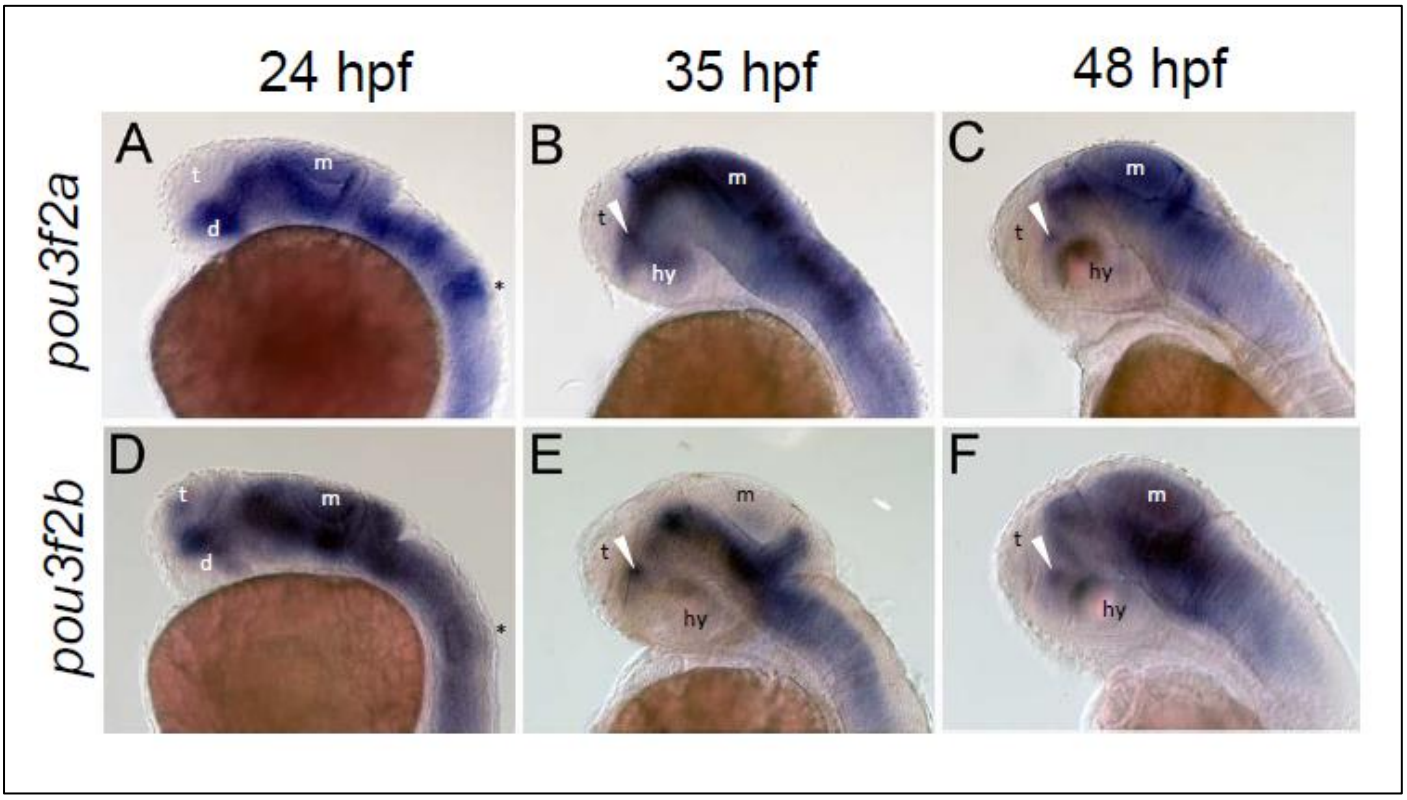


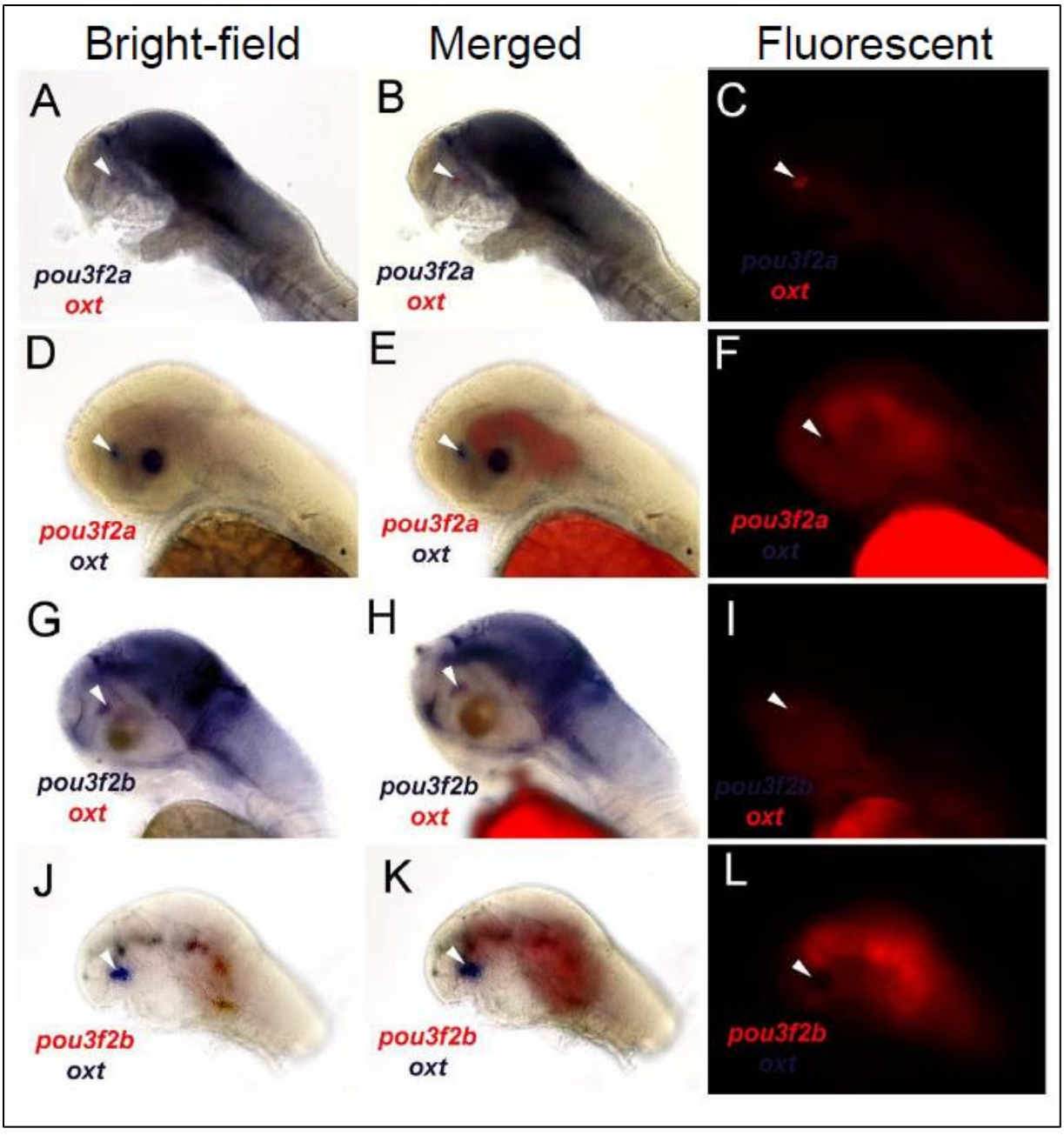
Figure S6: Developmental expression patterning of zebrafish *pou3f2a* and *pou3f2b.*
The expression patterns of *pou3f2a* and *pou3f2b* mRNAs were determined by Whole mount in situ hybridization (WISH) at 24, 35, and 48 hours post fertilization (hpf).
Lateral views, with anterior left, and dorsal up.



*Key: t, telencephalon; d, diencephalon; hy, hypothalamus; m, midbrain; *, marks the position of rhombomere r5.*

(**A, D**) At 24 hpf *pou3f2a* and *pou3f2b* are extensively expressed in the diencephalon, midbrain, and hindbrain. In the hindbrain, *pou3f2a* expression is concentrated in rhombomeres r1, r3 and r5.
(**B, E**) By 35 hpf, *pou3f2a* and *pou3f2b* expression becomes weaker in the anterior diencephalon along the telencephalic-diencephalic border, and in the posterior hypothalamus. Within the anterior diencephalic domain, strong *pou3f2a* expression is maintained in the neuroendocrine preoptic area (NPO), indicated with white arrowheads.
(**C, F**) By 48 hpf, *pou3f2a* and *pou3f2b* expression is seen in the diencephalon, the midbrain tegmentum and throughout the hindbrain. In the diencephalon, expression further narrows and strong expression of *pou3f2a* and *pou3f2b* remains restricted to a small area of the NPO.

Figure S7: Co-localization of *pou3f2a* and *pou3f2b* mRNAs with subsets of oxytocin cells.
Co-expression of *pou3f2a*, *pou3f2b* and *oxt* mRNA was determined by double label WISH in 48 hpf embryos. *Pou3f2a* and *pou3f2b* probes were labelled with digoxygenin (DIG) while oxytocin probes were labeled with fluorescein (FL). DIG and FL were detected using either BM Purple or Red Fast Stain.
The white arrowhead indicates the location of the *oxt* expression domain. (**A, D, G, J**) Bright field images (**C, F, I, L**) are the corresponding epifluorescence images, and (**B, E, H, K**) are the merged bright field and epifluorescence images.



(**A-C**) *pou3f2a* expression is visualized with purple staining and *oxt* expression is visualized with red staining. The eyes have been removed to better visualize the staining patterns. The purple *pou3f2a* staining quenches the red *oxt* stain, except in a small posterior region, of the *oxt* domain.
(**D-F**) The chromogenic staining is reversed such that *pou3f2a* expression is visualized with red staining and

oxt expression is visualized with purple staining. The purple *oxt* stain quenches the entire red *pou3f2a* stain in this region. Thus, *pou3f2a* appears to be co-expressed in a subset of oxytocin cells.
(**G-I**) *pou3f2b* expression is visualized with purple staining and *oxt* expression is visualized with red staining. The purple *pou3f2b* staining quenches the red *oxt* stain of the *oxt* domain, except for in a small posterior region as seen with *pou3f2a*.
(**J-L**) The chromogenic staining is reversed such that *pou3f2b* expression is visualized with red staining and *oxt* expression is visualized with purple staining. The eyes have been removed to better visualize the staining patterns. The purple *oxt* stain quenches the entire red *pou3f2b* stain in this region. Thus, similar to *pou3f2a*, *pou3f2b* appears to be co-expressed in a subset of oxytocin cells.

Figure S8: Knockdown of *pou3f2a* or *pou3f2b* has no effect on expression of *sim1a*

MO knockdown of *pou3f2a* or *pou3f2b* causes no obvious change in *sim1a* expression patterns. (**A-C**) Lateral views, 48 hpf embryos stained for *sim1a* expression by WISH. The location of the NPO is indicated with white arrowheads. (**A**) control MO injected embryos (**B**) *pou3f2a* MO injected embryo, (**C**) *pou3f2b* MO injected embryo.

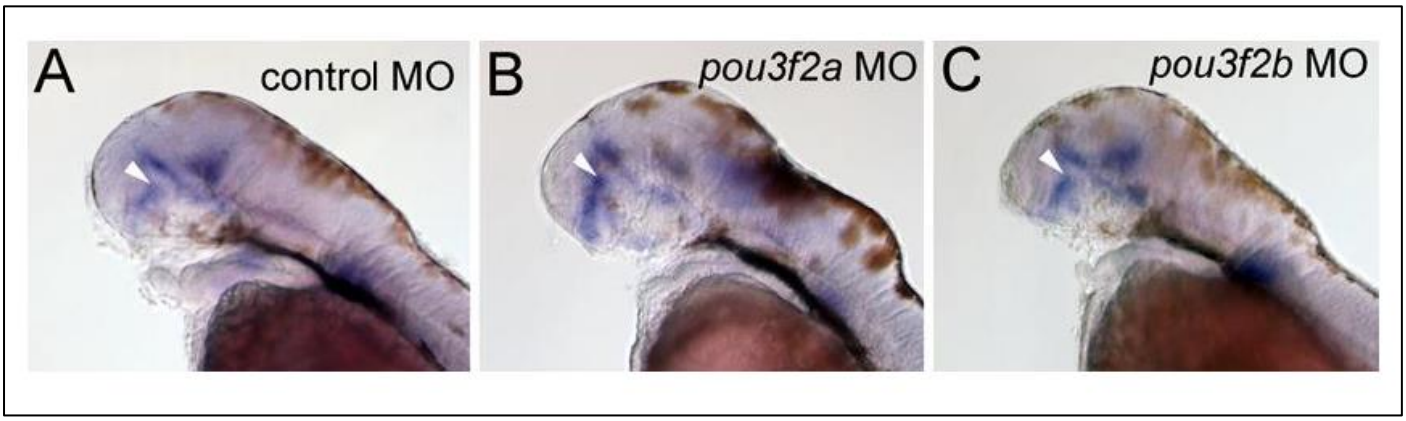
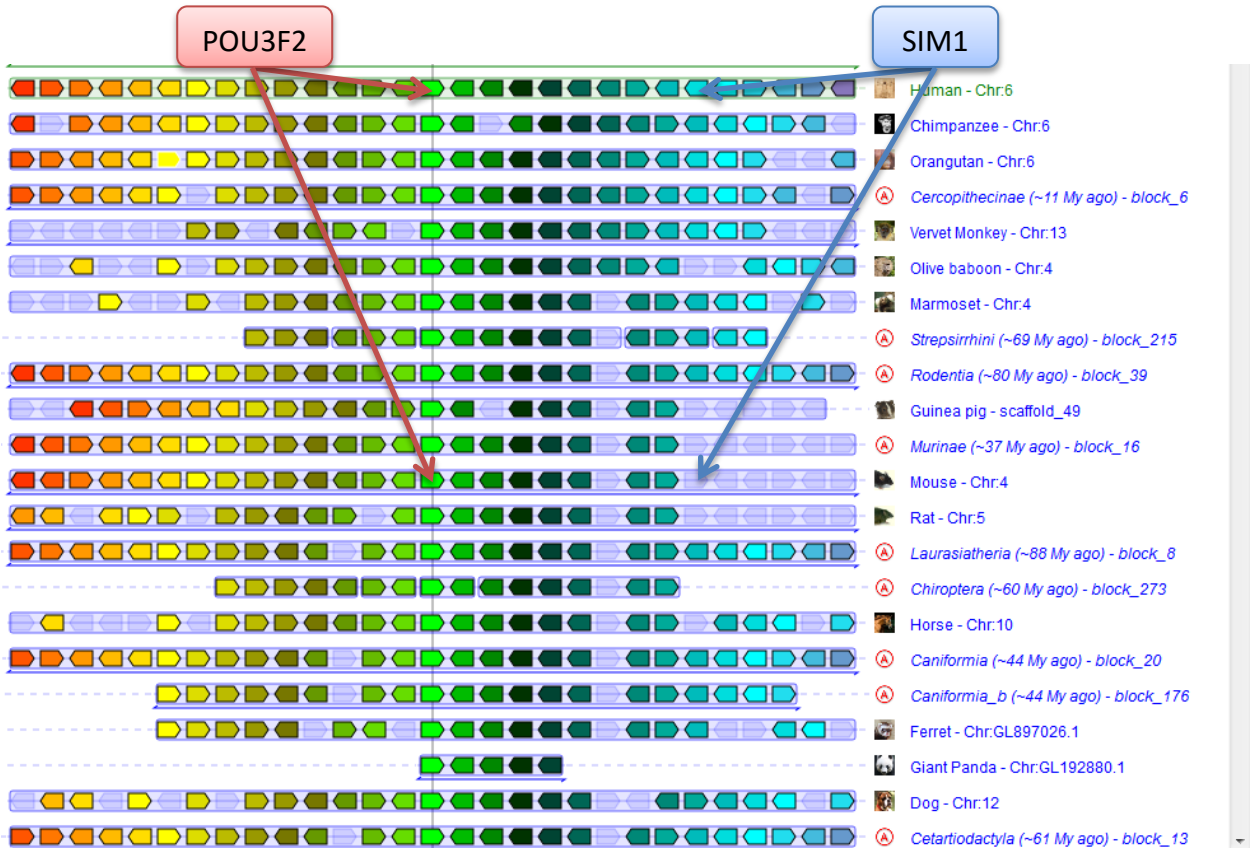


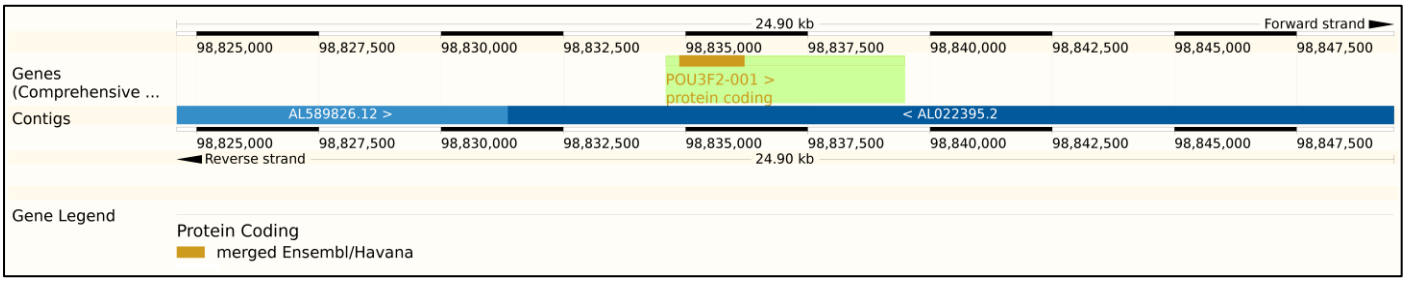
Figure S9: The evolutionary syntenic architecture of the 6q16.2 region shows that POU3F2 and SIM1 are not located on the same chromosome in a number of mammals.



Snapshot from Genomicus genome browser. The figure is centred on POU3F2 gene which is indicated with a red box and arrow. SIM1 location and orientation is indicated by blue box and arrow. The right column gives the name of the species. Faded boxes indicate that the genes are not syntenically linked in the respective specie. Mouse Pou3f2 and Sim1 are highlighted demonstrating that unlike in humans, these two genes are on different chromosomes.

Figure S10: ***POU3F2 is a single exon gene with a number of repeats and thus challenging to sequence***

A



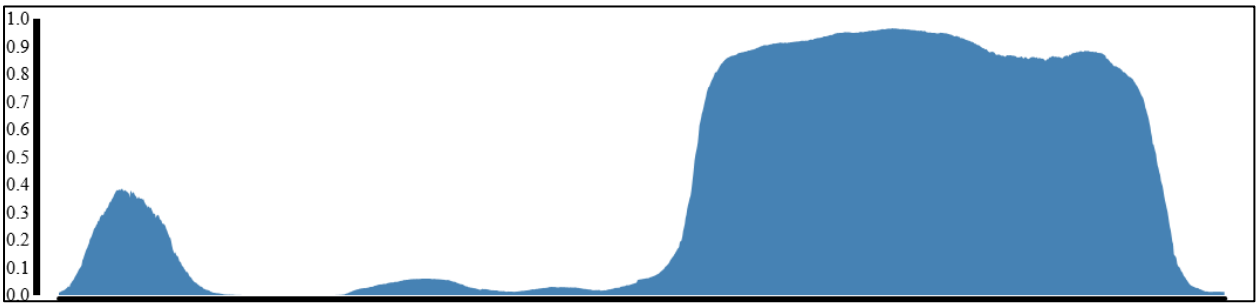
B

```
 241  GCGGCTCCTTTAACCGGAGCGCTCAGTCCGGCTCCGAGAGTCATGGCGACCGCAGCGTCT
      .........................................ATGGCGACCGCAGCGTCT
      .........................................-M--A--T--A--A--S-

 301  AACCACTACAGCCTGCTCACCTCCAGCGCCTCCATCGTGCACGCCGAGCCGCCCGGCGGC
  19  AACCACTACAGCCTGCTCACCTCCAGCGCCTCCATCGTGCACGCCGAGCCGCCCGGCGGC
   7  -N--H--Y--S--L--L--T--S--S--A--S--I--V--H--A--E--P--P--G--G-

 361  ATGCAGCAGGGCGCGGGGGGCTACCGCGAAGCGCAGAGCCTGGTGCAGGGCGACTACGGC
  79  ATGCAGCAGGGCGCGGGGGGCTACCGCGAAGCGCAGAGCCTGGTGCAGGGCGACTACGGC
  27  -M--Q--Q--G--A--G--G--Y--R--E--A--Q--S--L--V--Q--G--D--Y--G-

 421  GCTCTGCAGAGCAACGGACACCCGCTCAGCCACGCTCACCAGTGGATCACCGCGCTGTCC
 139  GCTCTGCAGAGCAACGGACACCCGCTCAGCCACGCTCACCAGTGGATCACCGCGCTGTCC
  47  -A--L--Q--S--N--G--H--P--L--S--H--A--H--Q--W--I--T--A--L--S-

 481  CACGGCGGCGGCGGCGGGGGCGGTGGCGGCGGCGGGGGGGGCGGGGGCGGCGGCGGGGGC
 199  CACGGCGGCGGCGGCGGGGGCGGTGGCGGCGGCGGGGGGGGCGGGGGCGGCGGCGGGGGC
  67  -H--G--G--G--G--G--G--G--G--G--G--G--G--G--G--G--G--G--G--G-

 541  GGCGGCGACGGCTCCCCGTGGTCCACCAGCCCCCTGGGCCAGCCGGACATCAAGCCCTCG
 259  GGCGGCGACGGCTCCCCGTGGTCCACCAGCCCCCTGGGCCAGCCGGACATCAAGCCCTCG
  87  -G--G--D--G--S--P--W--S--T--S--P--L--G--Q--P--D--I--K--P--S-

 601  GTGGTGGTGCAGCAGGGCGGCCGCGGAGACGAGCTGCACGGGCCAGGCGCCCTGCAGCAG
 319  GTGGTGGTGCAGCAGGGCGGCCGCGGAGACGAGCTGCACGGGCCAGGCGCCCTGCAGCAG
 107  -V--V--V--Q--Q--G--G--R--G--D--E--L--H--G--P--G--A--L--Q--Q-

 661  CAGCATCAGCAGCAGCAACAGCAACAGCAGCAGCAACAGCAGCAACAGCAGCAGCAGCAG
 379  CAGCATCAGCAGCAGCAACAGCAACAGCAGCAGCAACAGCAGCAACAGCAGCAGCAGCAG
 127  -Q--H--Q--Q--Q--Q--Q--Q--Q--Q--Q--Q--Q--Q--Q--Q--Q--Q--Q--Q-

 721  CAGCAACAGCGGCCGCCGCATCTGGTGCACCACGCCGCTAACCACCACCCGGGACCCGGG
 439  CAGCAACAGCGGCCGCCGCATCTGGTGCACCACGCCGCTAACCACCACCCGGGACCCGGG
 147  -Q--Q--Q--R--P--P--H--L--V--H--H--A--A--N--H--H--P--G--P--G-

 781  GCATGGCGGAGCGCGGCGGCTGCAGCGCACCTCCCACCCTCCATGGGAGCGTCCAACGGC
 499  GCATGGCGGAGCGCGGCGGCTGCAGCGCACCTCCCACCCTCCATGGGAGCGTCCAACGGC
 167  -A--W--R--S--A--A--A--A--A--H--L--P--P--S--M--G--A--S--N--G-

 841  GGCTTGCTCTACTCGCAGCCCAGCTTCACGGTGAACGGCATGCTGGGCGCCGGCGGGCAG
 559  GGCTTGCTCTACTCGCAGCCCAGCTTCACGGTGAACGGCATGCTGGGCGCCGGCGGGCAG
 187  -G--L--L--Y--S--Q--P--S--F--T--V--N--G--M--L--G--A--G--G--Q-

 901  CCGGCCGGTCTGCACCACCACGGCCTGCGGGACGCGCACGACGAGCCACACCATGCCGAC
 619  CCGGCCGGTCTGCACCACCACGGCCTGCGGGACGCGCACGACGAGCCACACCATGCCGAC
 207  -P--A--G--L--H--H--H--G--L--R--D--A--H--D--E--P--H--H--A--D-

 961  CACCACCCGCACCCGCACTCGCACCCACACCAGCAGCCGCCGCCCCCGCCGCCCCCGCAG
 679  CACCACCCGCACCCGCACTCGCACCCACACCAGCAGCCGCCGCCCCCGCCGCCCCCGCAG
 227  -H--H--P--H--P--H--S--H--P--H--Q--Q--P--P--P--P--P--P--P--Q-

1021  GGTCCGCCTGGCCACCCAGGCGCGCACCACGACCCGCACTCGGACGAGGACACGCCGACC
 739  GGTCCGCCTGGCCACCCAGGCGCGCACCACGACCCGCACTCGGACGAGGACACGCCGACC
 247  -G--P--P--G--H--P--G--A--H--H--D--P--H--S--D--E--D--T--P--T-
```

```
1081 TCGGACGACCTGGAGCAGTTCGCCAAGCAGTTCAAGCAGCGGCGGATCAAACTGGGATTT
 799 TCGGACGACCTGGAGCAGTTCGCCAAGCAGTTCAAGCAGCGGCGGATCAAACTGGGATTT
 267 -S--D--D--L--E--Q--F--A--K--Q--F--K--Q--R--R--I--K--L--G--F-

1141 ACCCAAGCGGACGTGGGGCTGGCTCTGGGCACCCTGTATGGCAACGTGTTCTCGCAGACC
 859 ACCCAAGCGGACGTGGGGCTGGCTCTGGGCACCCTGTATGGCAACGTGTTCTCGCAGACC
 287 -T--Q--A--D--V--G--L--A--L--G--T--L--Y--G--N--V--F--S--Q--T-

1201 ACCATCTGCAGGTTTGAGGCCCTGCAGCTGAGCTTCAAGAACATGTGCAAGCTGAAGCCT
 919 ACCATCTGCAGGTTTGAGGCCCTGCAGCTGAGCTTCAAGAACATGTGCAAGCTGAAGCCT
 307 -T--I--C--R--F--E--A--L--Q--L--S--F--K--N--M--C--K--L--K--P-

1261 TTGTTGAACAAGTGGTTGGAGGAGGCGGACTCGTCCTCGGGCAGCCCCACGAGCATAGAC
 979 TTGTTGAACAAGTGGTTGGAGGAGGCGGACTCGTCCTCGGGCAGCCCCACGAGCATAGAC
 327 -L--L--N--K--W--L--E--E--A--D--S--S--S--G--S--P--T--S--I--D-

1321 AAGATCGCAGCGCAAGGGCGCAAGCGGAAAAAGCGGACCTCCATCGAGGTGAGCGTCAAG
1039 AAGATCGCAGCGCAAGGGCGCAAGCGGAAAAAGCGGACCTCCATCGAGGTGAGCGTCAAG
 347 -K--I--A--A--Q--G--R--K--R--K--K--R--T--S--I--E--V--S--V--K-

1381 GGGGCTCTGGAGAGCCATTTCCTCAAATGCCCCAAGCCCTCGGCCCAGGAGATCACCTCC
1099 GGGGCTCTGGAGAGCCATTTCCTCAAATGCCCCAAGCCCTCGGCCCAGGAGATCACCTCC
 367 -G--A--L--E--S--H--F--L--K--C--P--K--P--S--A--Q--E--I--T--S-

1441 CTCGCGGACAGCTTACAGCTGGAGAAGGAGGTGGTGAGAGTTTGGTTTTGTAACAGGAGA
1159 CTCGCGGACAGCTTACAGCTGGAGAAGGAGGTGGTGAGAGTTTGGTTTTGTAACAGGAGA
 387 -L--A--D--S--L--Q--L--E--K--E--V--V--R--V--W--F--C--N--R--R-

1501 CAGAAAGAGAAAAGGATGACCCCTCCCGGAGGGACTCTGCCGGGCGCCGAGGATGTGTAC
1219 CAGAAAGAGAAAAGGATGACCCCTCCCGGAGGGACTCTGCCGGGCGCCGAGGATGTGTAC
 407 -Q--K--E--K--R--M--T--P--P--G--G--T--L--P--G--A--E--D--V--Y-

1561 GGGGGGAGTAGGGACACTCCACCACACCACGGGGTGCAGACGCCCGTCCAGTGAACTCGA
1279 GGGGGGAGTAGGGACACTCCACCACACCACGGGGTGCAGACGCCCGTCCAGTGA......
 427 -G--G--S--R--D--T--P--P--H--H--G--V--Q--T--P--V--Q--*-......
```

C

D



A. *Snapshot from Ensembl genome browser from Gene Summary page for POU3F2 (ENSG00000184486) demonstrating that this gene has only one exon.*
B. *POU3F2 cDNA sequence. The long repeats in the DNA sequence are highlighted.*
C. *Domains of POU3F2 protein*
D. *ExAC coverage plot for POU3F2. X-axis denotes nucleotide positions (unnumbered) in POU3F2 exon and y-axis number of individuals with more than 30X coverage for the corresponding nucleotide position on the X-axis.*

Table S1: Monogenic obesity syndromes in the Leptin>melanocortin>SIM1 neuro-endocrine pathway.

Gene	Phenotype	OMIM	Reference
LEP	Morbid obesity due to leptin deficiency	614962	[5]
LEPR	Morbid obesity due to leptin receptor deficiency	614963	[6,7]
POMC	Obesity, adrenal insufficiency and red hair syndrome	609734	[8]
PCSK1	Obesity with impaired prohormone processing	600955	[9]
MC4R	Autosomal dominant obesity	601665	[10,11]
MRAP2	Susceptibility to obesity BMIQ18	615457	[12]
SIM1	Severe obesity	601665	[13–15]

Table S2: List of genes within the deletion in Family 1 and properties of their variation.

Table based on data accessed from http://exac.broadinstitute.org/ on 15th September 2015.

Gene	Number of potentially truncating variants	Combined allele frequency of truncating variants	HI index	Associated disease (OMIM)	Inheritance pattern of the disease
POU3F2	0	0	15.18	None	Not applicable
FBXL4	11	$3.1x10^{-4}$	10.28	Encephalomyopathic type mitochondrial DNA depletion syndrome (605654)	Autosomal recessive
FAXC	2	$1.7x10^{-5}$	21.43	None	Not applicable
COQ3	19	$8.2x10^{-4}$	40.90	None	Not applicable
PNISR	11	$2.2\ x10^{-4}$	9.31	None	Not applicable
USP45	57	$1.9X10^{-2}$	45.26	None	Not applicable
TSTD3	No data	No data	No data	None	Not applicable
CCNC	10	$3.8x10^{-1}$	2.44	None	Not applicable
PRDM13	8	$8.1x10^{5}$	56.20	None	Not applicable

Table S3: Extent of 6q16 deletion in all the families studied in this project.

Family #	Start co-ordinate (Hg19)	End co-ordinate (Hg19)	Size (Mb)
1	chr6:99218535	chr6:100260996	1.04
2	chr6:98583798	chr6:99582590	0.99
3	chr6:98981807	chr6:100377741	1.39
4	chr6:98981812	chr6:100261128	1.28
5	chr6:99179231	chr6:100641841	1.46
6	chr6:99218553	chr6:100319568	1.10

Table S4: Amino acid identities in the amino terminal region of POU3 family proteins.

Outside of the homeodomain there is 59% identity between Pou3f2a and Pou3f2b, and 50% identity between POU3F2 and Pou3f2a or Pou3f2b. The amino acid identity among Pou3f2a/b and other members of the POU family is reduced in comparison.

POU family protein	Pou3f2a	Pou3f2b
zf_Pou3f2b	59%	-
Hu_POU3F1	35%	30%
Hu_POU3F2	50%	50%
Hu_POU3F3	42%	28%
Hu_POU3F4	47%	45%

Table S5: pou3f2a/b MO sequences.

pou3f2a 5’	AGTGGTTGGACGCCGCGGTCGCCAT
control-pou3f2a 5’	AGTcGTTGcACcCCGCcGTCcCCAT
pou3f2b 5'	GATTGGATGCTGTAGTCGCCATGAC
control-pou3f2b 5’	GAaTGcATGCTGTAGTCGCgATcAC

The underlined nucleotides mark the AUG translation initiation site, while the lower case letters in the control MO indicates the bases that are mis-matched.

Table S6: Primers and reaction conditions for arnt2^{hi2639cTg} genotyping

Hi2639_3E01 (F1):	5’ ATA CTG AGG GTG AAC GCA GAC G 3’
Hi2639_3E02 (R1):	5’ TCG CTT CTC GCT TCT GTT CG 3’
arnt2 Endogenous:	5’ AAA TCG CAT CCA AGC ATC GC 3’

Embryos were genotyped for the arnt2^{hi2639cTg} allele by tail clip. DNA was prepared by alkaline lysis (25 mM NaOH, 0.2 mM disodium EDTA) for 1 hour at 95° C, followed by neutralization with an equal amount of 40 mM Tris-HCl; pH5. Amplification conditions were as per ZIRC PCR protocol (http://zebrafish.org/zirc/fish/pdf/pcr/hi2639cTg.pdf), with the addition of a third primer to identify the endogenous allele. The PCR reaction was performed using 1X PCR Master Mix (Thermo Scientific), 0.4 uM F1 primer, 0.2 uM R1 primer, 0.2 uM Endogenous primer, and 4 µl of tail clip DNA, in a 25 µl reaction solution. The *arnt2* PCR program was run according to the following protocol: 94°C, 3 minutes; 94°C, 30 seconds; 62°C, 40 seconds; 72°C, 30 seconds; repeat steps 2-4, 34 cycles, 72°C, 5 minute extention. The entire reaction was run on a 2% agarose gel, producing a mutant band (arnt2^{hi2639cTg} = 342 bp), a wild-type band (wild-type = 237 bp) or both for the heterozygous genotype.

Supplementary References

1. Kang, H. J. *et al.* Spatio-temporal transcriptome of the human brain. *Nature* **478,** 483–489 (2011).
2. Hawrylycz, M. *et al.* in *Springer Handbook of Bio-/Neuroinformatics* (ed. Kasabov, N.) 1111–1126 (Springer Berlin Heidelberg, 2014).
3. Sugitani, Y. *et al.* Brn-1 and Brn-2 share crucial roles in the production and positioning of mouse neocortical neurons. *Genes Dev.* **16,** 1760–1765 (2002).
4. Belot, A. *et al.* Protein Kinase Cδ Deficiency Causes Mendelian Systemic Lupus Erythematosus With B Cell-Defective Apoptosis and Hyperproliferation. *Arthritis Rheum.* **65,** 2161–2171 (2013).
5. Farooqi, I. S. *et al.* Partial leptin deficiency and human adiposity. *Nature* **414,** 34–35 (2001).
6. Clement, K. *et al.* A mutation in the human leptin receptor gene causes obesity and pituitary dysfunction. *Nature* **392,** 398–401 (1998).
7. Farooqi, I. S. *et al.* Clinical and Molecular Genetic Spectrum of Congenital Deficiency of the Leptin Receptor. *N. Engl. J. Med.* **356,** 237–247 (2007).
8. Krude, H. *et al.* Severe early-onset obesity, adrenal insufficiency and red hair pigmentation caused by POMC mutations in humans. *Nat. Genet.* **19,** 155–157 (1998).
9. Jackson, R. S. *et al.* Obesity and impaired prohormone processing associated with mutations in the human prohormone convertase 1 gene. *Nat. Genet.* **16,** 303–306 (1997).
10. Vaisse, C., Clement, K., Guy-Grand, B. & Froguel, P. A frameshift mutation in human MC4R is associated with a dominant form of obesity. *Nat. Genet.* **20,** 113–114 (1998).
11. Farooqi, I. S. *et al.* Clinical Spectrum of Obesity and Mutations in the Melanocortin 4 Receptor Gene. *N. Engl. J. Med.* **348,** 1085–1095 (2003).
12. Asai, M. *et al.* Loss of Function of the Melanocortin 2 Receptor Accessory Protein 2 Is Associated with Mammalian Obesity. *Science* **341,** 275–278 (2013).

13. Holder, J. L., Butte, N. F. & Zinn, A. R. Profound obesity associated with a balanced translocation that disrupts the SIM1 gene. *Hum. Mol. Genet.* **9,** 101–108 (2000).
14. Bonnefond, A. *et al.* Loss-of-function mutations in SIM1 contribute to obesity and Prader-Willi–like features. *J. Clin. Invest.* **123,** 3037–3041 (2013).
15. Ramachandrappa, S. *et al.* Rare variants in single-minded 1 (SIM1) are associated with severe obesity. *J. Clin. Invest.* **123,** 3042–3050 (2013).