Rare variants in single-minded 1 (*SIM1*) are associated with severe obesity

Shwetha Ramachandrappa, … , Murray L. Whitelaw, I. Sadaf Farooqi

Research Article   Genetics

Single-minded 1 (SIM1) is a basic helix-loop-helix transcription factor involved in the development and function of the paraventricular nucleus of the hypothalamus. Obesity has been reported in *Sim1* haploinsufficient mice and in a patient with a balanced translocation disrupting *SIM1*. We sequenced the coding region of *SIM1* in 2,100 patients with severe, early onset obesity and in 1,680 controls. Thirteen different heterozygous variants in *SIM1* were identified in 28 unrelated severely obese patients. Nine of the 13 variants significantly reduced the ability of SIM1 to activate a SIM1-responsive reporter gene when studied in stably transfected cells coexpressing the heterodimeric partners of SIM1 (ARNT or ARNT2). *SIM1* variants with reduced activity cosegregated with obesity in extended family studies with variable penetrance. We studied the phenotype of patients carrying variants that exhibited reduced activity in vitro. Variant carriers exhibited increased ad libitum food intake at a test meal, normal basal metabolic rate, and evidence of autonomic dysfunction. Eleven of the 13 probands had evidence of a neurobehavioral phenotype. The phenotypic similarities between patients with SIM1 deficiency and melanocortin 4 receptor (MC4R) deficiency suggest that some of the effects of SIM1 deficiency on […]

Find the latest version:

http://jci.me/68016/pdf

Rare variants in single-minded 1 (*SIM1*) are associated with severe obesity

Shwetha Ramachandrappa,[1] Anne Raimondo,[2] Anna M.G. Cali,[1] Julia M. Keogh,[1] Elana Henning,[1] Sadia Saeed,[1] Amanda Thompson,[3] Sumedha Garg,[1] Elena G. Bochukova,[1] Soren Brage,[4] Victoria Trowse,[1] Eleanor Wheeler,[3] Adrienne E. Sullivan,[2] Mehul Dattani,[5] Peter E. Clayton,[6] Vippan Datta,[7] John B. Bruning,[2] Nick J. Wareham,[4] Stephen O'Rahilly,[1] Daniel J. Peet,[2] Ines Barroso,[1,3] Murray L. Whitelaw,[2] and I. Sadaf Farooqi[1]

[1]University of Cambridge Metabolic Research Laboratories and NIHR Cambridge Biomedical Research Centre, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, United Kingdom. [2]Discipline of Biochemistry, School of Molecular and Biomedical Science and Australian Research Council Special Research Centre for the Molecular Genetics of Development, University of Adelaide, Adelaide, Australia. [3]Wellcome Trust Sanger Institute, Cambridge, United Kingdom. [4]MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, United Kingdom. [5]Clinical and Molecular Genetics Unit, University College London Institute of Child Health, London, United Kingdom. [6]Manchester Academic Health Sciences Centre, Royal Manchester Children's Hospital, Manchester, United Kingdom. [7]Norfolk and Norwich University Hospital NHS Foundation Trust, Norwich, United Kingdom.

Single-minded 1 (SIM1) is a basic helix-loop-helix transcription factor involved in the development and function of the paraventricular nucleus of the hypothalamus. Obesity has been reported in *Sim1* haploinsufficient mice and in a patient with a balanced translocation disrupting *SIM1*. We sequenced the coding region of *SIM1* in 2,100 patients with severe, early onset obesity and in 1,680 controls. Thirteen different heterozygous variants in *SIM1* were identified in 28 unrelated severely obese patients. Nine of the 13 variants significantly reduced the ability of SIM1 to activate a SIM1-responsive reporter gene when studied in stably transfected cells coexpressing the heterodimeric partners of SIM1 (ARNT or ARNT2). *SIM1* variants with reduced activity cosegregated with obesity in extended family studies with variable penetrance. We studied the phenotype of patients carrying variants that exhibited reduced activity in vitro. Variant carriers exhibited increased ad libitum food intake at a test meal, normal basal metabolic rate, and evidence of autonomic dysfunction. Eleven of the 13 probands had evidence of a neurobehavioral phenotype. The phenotypic similarities between patients with SIM1 deficiency and melanocortin 4 receptor (MC4R) deficiency suggest that some of the effects of SIM1 deficiency on energy homeostasis are mediated by altered melanocortin signaling.

Introduction

Genetic disruption of several molecules involved in the hypothalamic leptin/melanocortin pathway can cause obesity in rodents and humans (1). Single-minded 1 (SIM1) is a basic helix-loop-helix transcription factor involved in the development and function of the paraventricular nucleus (PVN) of the hypothalamus, a region that plays a critical role in energy homeostasis (2). *Sim1* homozygous null mice (*Sim1*[−/−]) die in the perinatal period and have abnormal hypothalamic architecture, lacking paraventricular, anterior periventricular, and supraoptic nuclei because of failure of terminal migration and differentiation of Sim1-expressing neurons (3). *Sim1* heterozygous mice (*Sim1*[+/−]) survive and develop severe obesity associated with increased food intake without any measurable deficit in energy expenditure, establishing a role for *Sim1* in energy homeostasis (3–5). Conditional postnatal deletion of *Sim1* in postmitotic neurons recapitulates the obesity phenotype seen in germline depletion of *Sim1* in mice, indicating that *Sim1* has a physiological role in energy homeostasis that is distinct from its role in hypothalamic development (6).

Several lines of evidence suggest that SIM1 may influence energy homeostasis by interacting with pathways involved in central melanocortin signaling. The obesity phenotype of *Sim1*[+/−] mice

closely resembles that of mice lacking the melanocortin 4 receptor (Mc4r) with increased linear growth and increased feeding efficiency in addition to an increase in food intake (3, 4, 6). However, in contrast to *Mc4r*-null mice, energy expenditure is not reduced in *Sim1*[+/−] mice, suggesting that the relationship between Sim1 and Mc4r is not a simple linear one (4, 6). Sim1 is expressed in the PVN, basomedial amygdala, anterior hypothalamus, and lateral hypothalamic area (5), regions that also have high concentrations of *Mc4r*-expressing neurons, and *Mc4r* mRNA levels in the PVN are significantly reduced in *Sim1*[+/−] mice (6). Administration of melanotan II, a potent melanocortin receptor agonist, results in a reduction in food intake in WT mice, whereas this response is blunted in *Sim1*[+/−] mice, suggesting that Sim1 lies downstream of Mc4r (4). Transgenic overexpression of human SIM1 via a BAC transgene in *Agouti* mice (which ectopically overexpress the melanocortin receptor antagonist Agouti and share many of the features of *Mc4r*-null mice) ameliorates their phenotype by normalizing their food intake, suggesting that increased SIM1 expression is able to compensate for impaired Mc4r signaling (7).

There is some evidence that common variants at the SIM1 locus are associated with BMI (8, 9). Also, patients with chromosomal deletions involving 6q14–q21, a region that encompasses several genes including *SIM1*, have been reported to develop early onset obesity and developmental delay, with some reports suggesting that there is a degree of phenotypic overlap with the Prader-Willi syndrome (PWS) (10). In 2000, Holder and colleagues reported a female patient with severe early onset obesity and a de novo bal-
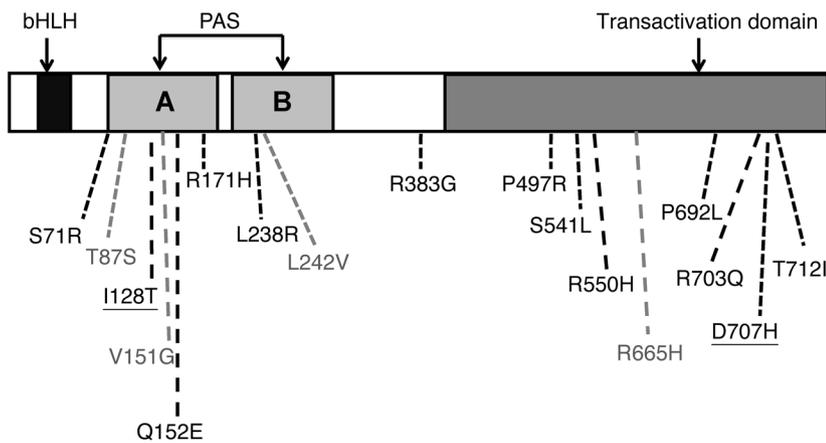
Figure 1
Identification of SIM1 variants in this study. Schematic of the domain structure of the human SIM1 protein. SIM1 variants identified in GOOS patients only (black), controls only (gray), and in both patients and controls (underlined) are shown.

anced translocation between chromosomes 1p22.1 and 6q16.2, which disrupted 1 allele of *SIM1* (11). We previously studied 277 patients with severe early onset obesity and identified a patient with a missense variant in SIM1 (I128T); however, in the absence of cosegregation in the immediate family, the pathogenicity of this variant was unclear (12). In order to comprehensively address whether variants in SIM1 contribute to obesity, we have now sequenced the *SIM1* gene in a large number of patients of mixed European descent with severe, early onset obesity recruited to the Genetics of Obesity Study (GOOS) (n = 2100) and a similar number of population-based controls from a large United Kingdom (UK) population-based study, the Ely study (n = 1680).

SIM1 belongs to the bHLH/PAS (basic helix-loop-helix/Per-Arnt-Sim) family of transcription factors, which are characterized by an N-terminal bHLH domain required for DNA binding and dimerization and a PAS domain that acts as a secondary dimerization interface (13, 14). Transcriptional control of target genes requires heterodimerization with another transcription factor, aryl-hydrocarbon receptor nuclear translocator (ARNT), or a homolog prevalent in the CNS, ARNT2 (15, 16). We went on to characterize the functional consequences of SIM1 variants found in patients using a luciferase reporter gene assay. Furthermore, we report the clinical phenotype of individuals carrying SIM1 variants with reduced activity and find several parallels with the phenotype of MC4R deficiency.

Results

Identification of SIM1 variants. The coding region of *SIM1* was sequenced in 2,100 unrelated patients with severe, early onset obesity recruited as part of GOOS and 1,680 unrelated population-based controls; both groups were of mixed European descent. No nonsense or frameshift variants were found in either group. Multiple nonsynonymous missense variants were found in obese individuals (28/2100) and in controls (14/1680) (Figure 1 and Supplemental Table 1; supplemental material available online with this article; doi:10.1172/JCI68016DS1). Several of the SIM1 variants found only in severely obese patients were predicted to be functionally deleterious based on SIFT (Sorting Intolerant From Tolerant) analysis and altered residues that were highly conserved across species (Supplemental Table 1), suggesting they may have functional consequences. Where possible, we sequenced the *SIM1* gene in available family members of severely obese probands. The majority of SIM1 variants found exclusively in patients cosegre-

gated with obesity in affected families with variable penetrance across and within families (Figure 2).

Three nonsynonymous variants were found in multiple obese cases, but also in controls in this study and in publicly available exome data: I128T (rs138546433), Q152E (rs140908824), and D707H (rs74726213) (Figure 1 and Supplemental Table 1). Notably, I128T has been reported at the same frequency in normal weight and obese populations in previous studies (9). The mean BMI of controls carrying these 3 variants was 27.8 ± 2.1 kg/m²; however, family members of controls were not available for genotyping, so the contribution of these variants to final body weight remains to be tested. Four variants were found only in controls; as 3 of these variants were predicted to be tolerated by SIFT analysis, these variants were not studied further (Supplemental Table 1).

Functional properties of mutant SIM1 proteins. In order to study the functional properties of the mutant SIM1 proteins found in patients, cDNAs expressing either WT or mutant myc-epitope–tagged SIM1 were incorporated into a single defined genetic locus of HEK293 Flp-In T-Rex cells, in which the expression of SIM1 was constitutively repressed by an upstream tetracycline repressor element. The addition of doxycycline induced expression of SIM1 protein of the expected approximately 85 kDa. Consistent levels of induction were seen with different clonal cell lines expressing the same protein (e.g., WT SIM1) and similar levels of expression were observed across the set of SIM1 mutant proteins (Figure 3A).

The bona fide DNA recognition sequence of SIM1 is unknown; however, it has been observed to bind the CNS midline enhancer element (CME) TACGTG, which is the target sequence of the drosophila homolog of SIM1(17). Multimerized CME has previously been shown to mediate SIM1-dependent transcriptional activation of an adenovirus major late promoter–driven firefly luciferase reporter gene (pML/6C-WT) in mammalian cells (15, 18). This reporter plasmid and expression vectors for either ARNT or ARNT2 were introduced into WT SIM1 inducible cells and luciferase activity measured after 16 hours treatment with doxycycline. Similar activity was observed with 2 independently derived WT SIM1 stable cell lines (data not shown), with coexpression with ARNT giving higher activity than coexpression of ARNT2 (data not shown).

To examine the functional consequences of the variants on the transcriptional activity of SIM1, we compared the ability of mutant SIM1 proteins to activate the pML/6C-WT reporter with WT SIM1 upon coexpression of either ARNT or ARNT2. While ARNT2 is the proposed in vivo partner of SIM1 (15), experiments

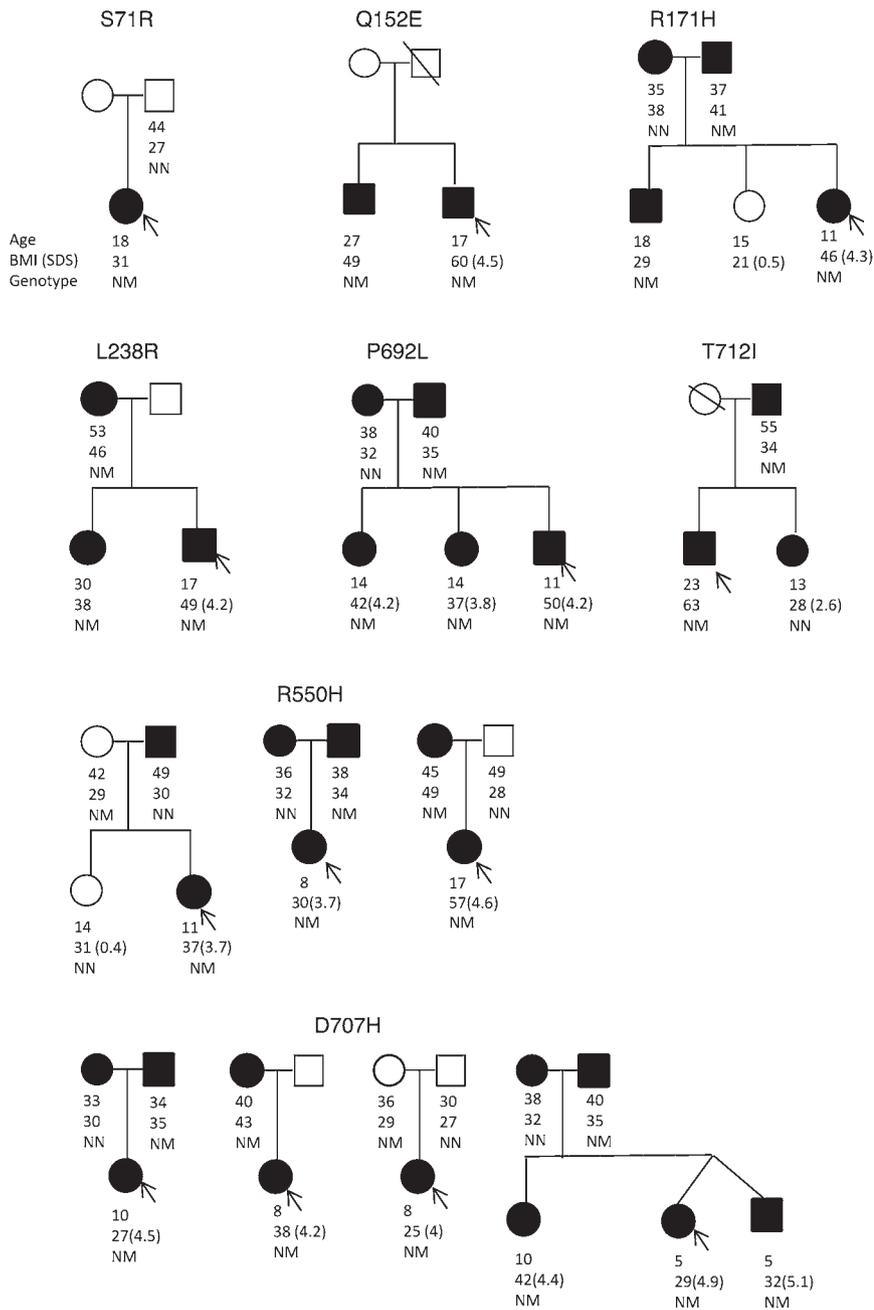**Figure 2**
Pedigrees of families of SIM1 variant carriers. Squares and circles represent male and female family members, respectively. Black symbols denote family members with obesity, arrows indicate probands. Below each individual, the age and BMI (SD score for children) is given along with genotype, where N represents a normal SIM1 allele and M represents a mutant allele.

might disrupt nuclear translocation by performing immunocytochemistry on cell lines expressing WT SIM1-myc and R383G SIM1-myc (Supplemental Figure 1). While both WT and mutant forms of SIM1 were predominantly nuclear, the R383G mutant had increased cytoplasmic staining compared with WT, suggesting less efficient nuclear uptake of this variant. However, this altered efficiency did not result in a decrease of activity in our reporter gene assays, suggesting the synthetic reporter system may underestimate the functional consequences of some variants.

Using the crystal structure for the N-terminal half of the CLOCK/BMAL heterodimer, we constructed a homology model to predict the likely structural features of the SIM1/ARNT2 bHLH/PAS-A heterodimer. This revealed the N-terminal amino acids that harbor variants, S71, I128, Q152, and R171, are likely to lie on the surface of the protein (Figure 3D). It is therefore possible that these amino acids contact accessory proteins important for SIM1 function. S71, I128, and Q152 are all situated in unstructured loops (Figure 3, E and F), making it difficult to predict how these variants might influence SIM1 structure. In contrast, R171 is situated in a helix that is in close proximity to other helices that make up PAS-A of SIM1 and is within range for intramolecular hydrogen bonding with both the peptide backbone and other residues (Figure 3E). Thus, the least active variant in the reporter assay, R171H, may contain a destabilized PAS A

helix, which could weaken dimerization and/or decrease overall protein stability. As crystal structures show very different PAS B interactions for each of HIF/Arnt PAS B (20), CLOCK/BMAL PAS B (21), and the Per PAS A.PAS B homodimer (22), we have not attempted modeling beyond PAS-A. Modeling was not possible for the C terminus, as there are no structures of this region for any member of the protein family.

Phenotype of SIM1 variant carriers. Patients harboring variants with reduced activity in *SIM1* and their relatives were invited for clinical studies including measurement of body composition, energy intake, energy expenditure, and endocrine function (Supplemental Table 1). Given the evidence suggesting that Sim1 is involved in melanocortin signaling in rodents, we compared the data from
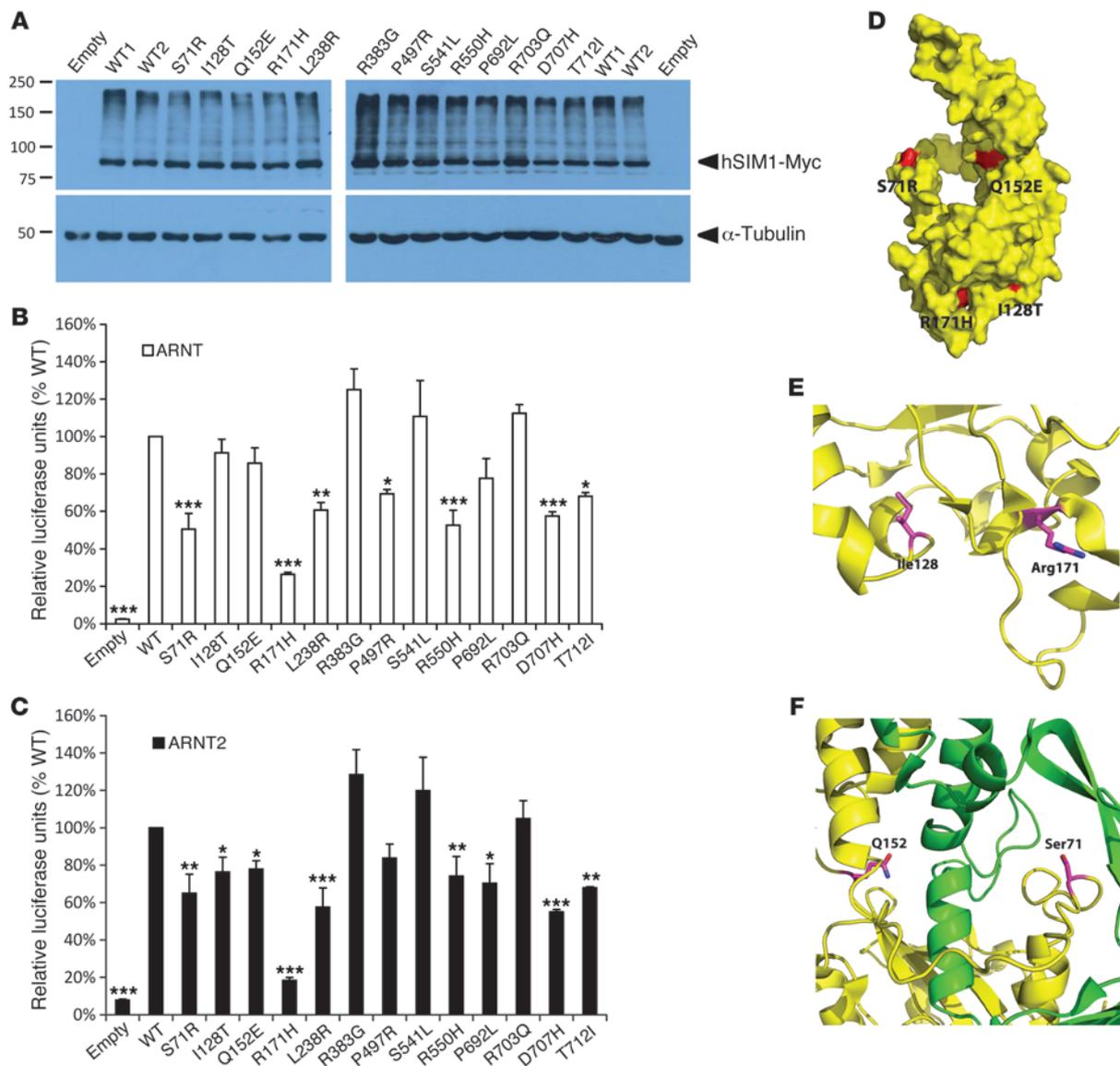
with ARNT were also performed to independently verify the trend of changing activities across the set of SIM1 mutants. Nine of the variants showed significant reduction of transcriptional activity in this assay when partnered with ARNT2, with R171H resulting in a profound reduction of activity with both partner proteins (*P* < 0.001) (Figure 3). Activities of 3 mutant proteins, R383G, S541L, and R703Q, were indistinguishable from WT in these assays. Transcriptional activity in complex with ARNT2 was more severely affected than that with ARNT, although the general trend of comparative activities was very similar.

Interestingly, R383 forms part of a novel nuclear localization signal (NLS) elucidated in SIM1 and its paralog, SIM2 (19). We tested to determine whether loss of arginine at this position

Figure 3
Variants in SIM1 significantly reduce transcriptional activation. (**A**) Western blot analysis of Myc-tagged SIM1 expression in whole-cell extracts from doxycycline-induced 293 Flp-In T-Rex cells. (**B** and **C**) Two independently derived WT SIM1 293 Flp-In T-Rex stable cell lines, 13 different SIM1 mutant 293 Flp-In T-Rex stable cell lines, and 1 empty vector cell line were transiently transfected with SIM1 luciferase reporter plasmid, Renilla control plasmid, and either ARNT (**B**) or ARNT2 (**C**) expression plasmid. Cells were then stimulated with doxycycline and analyzed for luciferase activity. Bars represent the mean Firefly/Renilla value expressed relative to WT, which has been normalized to 100%. Results shown are the average of at least 3 experiments performed in triplicate + SEM, expressed relative to the 2 WT lines, which have been combined into a single average value and normalized to 100%. Log values generated for each mutant were compared with WT. $*P < 0.05$; $**P < 0.01$; $***P < 0.001$. (**D**) Homology model of the SIM1 ARNT2 PAS-A heterodimer. SIM1 is shown in yellow and ARNT2 is shown in green, with amino acids that contain point variants colored red. (**E** and **F**) Ribbon diagrams of the SIM1 ARNT2 PAS-A heterodimer, with amino acids that contain point variants depicted as sticks colored pink.

patients carrying SIM1 variants that exhibited reduced activity in vitro to that obtained from patients with MC4R deficiency as well as comparably obese subjects with WT MC4R and SIM1 sequences, studied using the same methods.

Anthropometry and body composition. All SIM1 probands had increased body weight, deviating from predicted percentiles within the first 2 years of life. The mean BMI SD score of SIM1 variant carriers was 4.4 ± 0.4. In childhood, linear growth was nor-mal; mean height SD scores for prepubertal children with SIM1 variants (0.96 ± 0.35) were comparable to those of equally obese children without SIM1 variants (Figure 4). Final height in SIM1-deficient adults was not increased (Figure 4). The lack of acceler-ated linear growth in patients with SIM1 mutations contrasts with the phenotype of the patient reported by Holder et al. (11) and the phenotype of Sim1 heterozygous mice. This difference may be explained by partial versus complete loss of SIM1 activity.
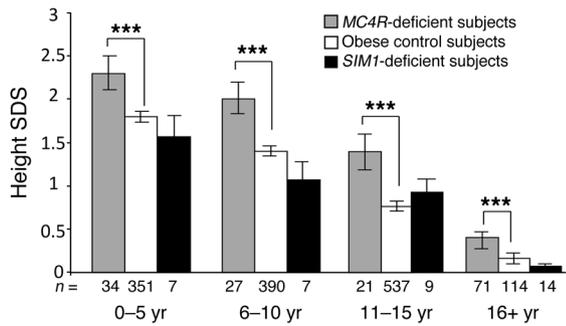
Figure 4
Measures of height SD scores of *MC4R*-deficient patients, SIM1 variant carriers, and obese control subjects. Values are mean ± SEM. ***$P < 0.001$.

We measured the body composition of SIM1 variant carriers by dual energy X-ray absorptiometry (DEXA). The mean percentage of body fat among SIM1 variant carriers was high at 45.8 ± 3.2 percent (normal range, 15 to 25 percent; $P < 0.001$) and comparable to that seen in patients with heterozygous MC4R variants (46.5 ± 2.2).

Energy intake and energy expenditure. All subjects had a history of increased food-seeking behavior in childhood, which continued into later life in the subset of adult subjects. The energy consumed by carriers of *SIM1* variants at an 18MJ ad libitum test meal given after an overnight fast was significantly greater than that of normal weight control subjects, after adjustment for lean body mass (Figure 5A). In this test, energy intake was comparable with MC4R-deficient patients. Basal metabolic rate (BMR) measured by indirect calorimetry was identical to that predicted based on the basis of age, sex, and body composition in children and adults (Figure 5B). The respiratory quotient (ratio of carbohydrate to fat oxidation) was significantly increased in SIM1 variant carriers compared with obese controls, but comparable to that seen in *MC4R*-deficient patients (Figure 5C).

Metabolic and endocrine function. Most of the subjects with SIM1 deficiency were euglycemic, although 1 adult had developed impaired glucose tolerance. All patients were hyperinsulinemic, but to a degree that was consistent with the degree of obesity (Figure 6). This contrasts with the severe disproportionate hyperinsulinemia we have previously reported in MC4R deficiency (23). Basal and thyrotrophin-releasing hormone-stimulated thyroid-stimulating hormone (TSH) levels were within normal limits, as were circulating levels of free thyroxine (data not shown).

Autonomic dysfunction. We studied blood pressure after an overnight fast in the recumbent position in SIM1 variant carriers, *MC4R*-deficient patients, and obese controls (Figure 7). Mean systolic blood pressure was lower in SIM1 variant carriers than in obese con-

trols; diastolic blood pressure measurements were not statistically different from those of controls (Figure 7, A and B). These observations are consistent with our previous observations in *MC4R*-deficient patients (24). To further examine autonomic nervous system function in SIM1 variant carriers, we measured heart rate variability at beat-to-beat resolution over 24 hours using a portable monitor. During sleep, there were no significant group differences in average heart rate, the high-frequency (HF) component of heart rate variability, and the root mean square of successive differences (RMSSD) between adjacent R-R intervals (predominantly measures of parasympathetic activation) (data not shown). On waking, the heart rate increased in all 3 groups, but the magnitude of the increase was attenuated in SIM1 variant carriers and in *MC4R*-deficient patients compared with obese controls (Figure 7C). These findings are consistent with impaired sympathetic tone in these subjects.

Neurobehavioral abnormalities. Eleven of the 13 probands who agreed to further studies had a degree of cognitive deficit as reported by physicians and family members. All of these patients carried variants that were associated with reduced activity in vitro (Supplemental Table 1). A spectrum of behaviors was seen, including impaired concentration, memory deficit, emotional lability, and in some cases, autistic spectrum behavior. Some of these features overlap with those reported in PWS, but of note, none of the patients had a reported history of hypotonia or feeding difficulties in early life that are hallmarks of PWS. Further behavioral testing was not possible as part of this study.

Discussion

In the present study, we sequenced the coding region of the transcription factor SIM1 in 2,100 patients of mixed European descent
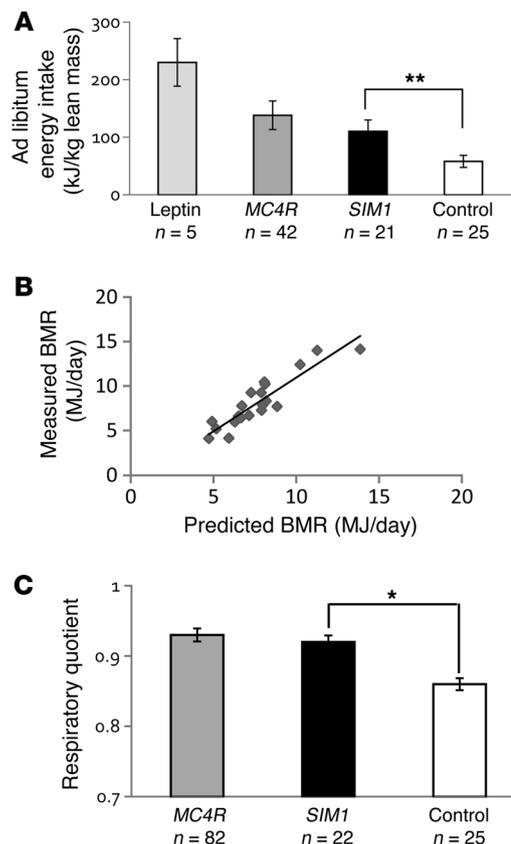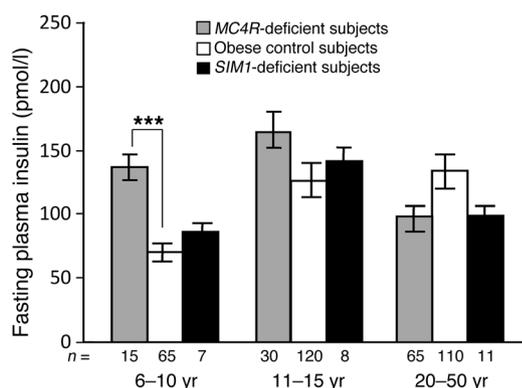

Figure 5
Phenotype of SIM1 variant carriers. (**A**) Ad libitum energy intake at an 18MJ test meal. Intake is expressed as kj/kg lean mass in *SIM1* variant carriers ($n = 23$) as compared with leptin-deficient patients ($n = 5$), *MC4R*-deficient patients ($n = 42$), and control subjects ($n = 25$). Values are mean ± SEM. **$P < 0.01$. (**B**) Measured versus predicted BMR adjusted for kilograms of fat free mass in SIM1 variant carriers. (**C**) Respiratory quotient as measured by indirect calorimetry in *SIM1* variant carriers ($n = 25$), *MC4R*-deficient patients ($n = 82$), and obese control subjects ($n = 25$). Values are mean ± SEM. *$P < 0.05$.

Figure 6
Fasting plasma insulin in *MC4R*-deficient patients, SIM1 variant carriers, and obese control subjects. ***P < 0.001. Values are mean ± SEM.

variants decreased SIM1 function, with R171H having the most profound effect on transcriptional activation. These findings suggest that the PAS-A and PAS-B domains in human SIM1 play a critical structural or functional role or roles that merit further investigation. Interestingly, I128T and Q152E were also found in controls, some, but not all, of whom were obese. The contribution of these variants to final body weight remains to be tested.

The remaining variants were found in the C terminus which is thought to be responsible for transcriptional regulation. R550H, D707H, and T712I were the most severely impaired, suggesting these are crucial residues in transcriptional regulation. As with other bHLH/PAS proteins (28, 29), there is evidence that the C terminus functions primarily as a transactivation or transrepression domain. By analogy, the closely related bHLH/PAS transcription factor SIM2 is known to regulate transcription differentially in a context-dependent manner (30, 31). SIM1 has variously been shown to activate and repress transcription in cell-based reporter assays. In addition to the reporter assays shown in Figure 3B, Ema et al. (32) showed that Sim1 repressed the ability of a Gal4-ARNT fusion protein to activate a Gal4 reporter gene, while deletion of the C terminus of Sim1 abolished this repression. The altered activities of our C-terminal SIM1 variants confirm the need to embark on experiments aimed at discovering SIM1 interacting proteins and direct target genes.

The activity of 3 of the variant SIM1 proteins (R383G, S541L, R703Q), was indistinguishable from that of WT in assays with

with severe, early onset obesity from the GOOS cohort and 1680 patients from a UK population-based control cohort. Cosegregation studies in families of SIM1 variant carriers indicated a dominant form of inheritance with variable penetrance. Some rare variants found in patients were also found in controls (e.g., I128T, Q152E, D707H). These variants are unlikely to be sufficient to explain the severe obesity in patients in this study, although they may be contributors to overweight/obesity in individuals with a susceptible genetic background. The SIM1-containing locus has not emerged from genome-wide association studies of BMI and 2 large studies in European populations have failed to find a significant association between common variants in SIM1 and increased BMI (25).

To explore the functional consequences of these variants, we generated stable cell lines expressing the mutant SIM1 proteins and tested their ability to activate the transcription of a SIM1-responsive reporter gene. Of the 13 variants examined, 9 had significantly reduced activity in this assay when paired with the proposed in vivo partner, ARNT2. A tenth variant, P497R, showed reduced activity but reached significance at *P* < 0.05 only with ARNT. No variants were found in the bHLH domain of the protein in this study, which mediates DNA binding and dimerization with ARNT and ARNT2. We identified 1 missense variant, S71R, in the sequence between the bHLH domain and the PAS region. This variant has weak activity, implying that this stretch of residues has a hitherto unrecognized functional role.
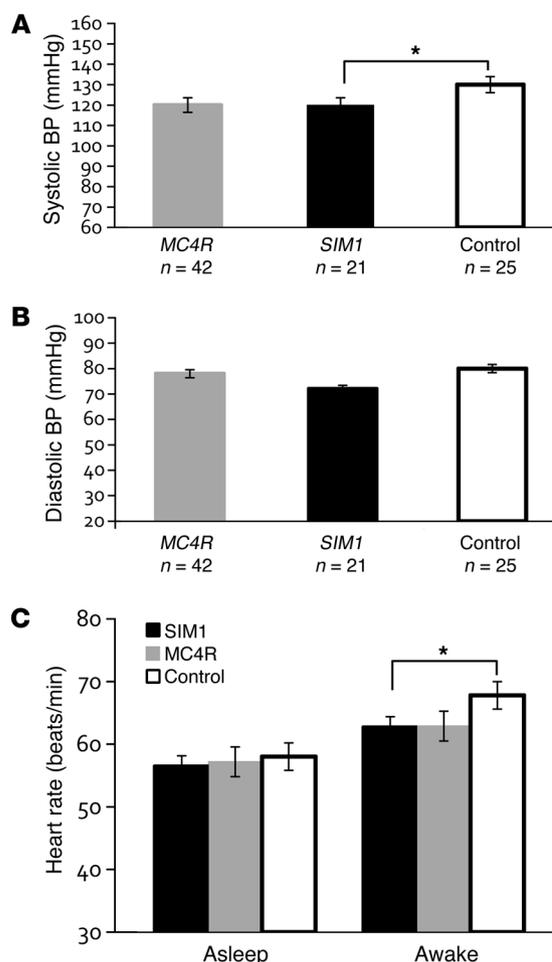
Three variants (I128T, Q152E, and R171H) were identified in the PAS-A domain and 1 (L238R) in the PAS-B domain. The PAS domains form conserved 3D folds and have been shown to perform multiple functions in bHLH/PAS proteins, including ligand binding (in the aryl hydrocarbon receptor), stabilization of heterodimerization (in CLOCK/BMAL; ref. 21), and interaction with either transcription coactivators (HIF-1A/ARNT; ref. 26) or repressive proteins (Cry interaction with Clock; ref. 27). All 4 PAS region

Figure 7
Autonomic function in SIM1 variant carriers. Measures of systolic (**A**) and diastolic (**B**) blood pressure are shown for SIM1 variant carriers (*n* = 23), MC4R deficient patients (*n* = 42) and obese control subjects (*n* = 25). (**C**) Heart rate during sleep and in the awake state is shown for SIM1 variant carriers (*n* = 23), *MC4R*-deficient patients (*n* = 42), and obese control subjects (*n* = 25). Values are mean ± SEM.

ARNT or ARNT2. It is possible that these variants are not pathogenic — there were no family data or detailed phenotypic data available for these patients, as they did not consent to further investigation (Supplemental Table 1). Alternatively, these variants may have a selective effect on aspects of SIM1 function that were not evaluated in our system. It is important to consider that, while this assay exclusively examined the transactivation properties of SIM1, the physiological relevance of this function remains to be established.

We have linked the development of severe early onset obesity with naturally occurring rare variants in SIM1 that cause reduced activity of this transcription factor. We report variants that reduce activity to varying degrees, from severely damaging variants that are likely to be major drivers of severe obesity (S71R, R171H, L238R, P497R, R550H, T712I) through moderately damaging variants that in some cases demonstrate variable penetrance (D707H) to variants that appear to result in a small reduction in activity whose precise contribution to obesity in carriers remains to be established (I128T, Q152E, R383G, S541L, P692L, R703Q). The reporter gene assay used in these studies can identify variants that are clearly deficient in activity, but may not be able to differentiate variants with mild loss of activity compared with WT SIM1. Given their location in the structure of SIM1, it is likely that some SIM1 variants may reduce function of the protein through a number of different mechanisms not tested here, including protein turnover, nuclear translocation, dimerization, DNA binding and binding to as-yet-unknown transcription cofactors.

Cumulatively, our genetic and functional data, together with the findings of Bonnefond and colleagues (33), provide direct evidence for a role for SIM1 in the control of energy homeostasis in humans. SIM1 is expressed in several regions of the brain, including the PVN of the hypothalamus, which is involved in the regulation of food intake. Mouse models have shown Sim1 to play both a developmental and a postnatal physiological role in this nucleus (2–4, 6, 34), with evidence divided as to whether 1 functional copy of *Sim1* is sufficient for its developmental function (4, 35). Consistent with postnatal deletion of *Sim1* causing hyperphagia (6), stereotaxic injection of an adenovirus expressing shRNA directed against *Sim1* into the PVN has shown that reduction of *Sim1* expression in a subset of neurons is sufficient to affect food intake (36)

Ad libitum energy intake was increased in patients with SIM1 variants with no evidence of a defect in basal energy expenditure. SIM1 variant carriers had lower blood pressure measurements, showed an attenuated increase in heart rate on waking, and had increased respiratory quotients when compared with obese controls. These later phenotypic characteristics were shared with *MC4R*-deficient patients and are indicative of underlying autonomic dysfunction that is seen in animals and humans with reduced melanocortinergic tone (24). These findings are consistent with a body of evidence that suggests a role for SIM1 in central melanocortin signaling.

Our results suggest that SIM1, or an-as-yet-unidentified transcriptional target of SIM1, forms part of the melanocortin signaling pathway. The melanocortin agonist MTII and leptin administration have been shown to increase *Sim1* gene expression in rodents, suggesting that *Sim1* is upregulated as a consequence of Mc4r activation(4). While the interaction between SIM1 and MC4R may operate at a number of levels, rescue of *Mc4r* expression selectively in *Sim1*-expressing neurons in *Mc4r*-knockout mice normalizes food intake and improves the body weight phenotype (37). This implies that *Mc4r* and *Sim1* are coexpressed in a sub-

set of neurons and that *Mc4r* in these neurons is critical to the central melanocortin pathway, which influences body weight by modulating food intake independently of energy expenditure (37). However, targeted ablation of *Sim1*-expressing neurons with diphtheria toxin resulted in obesity resultant from both hyperphagia and reduced energy expenditure, suggesting additional *Sim1* independent roles exist in at least some *Sim1*-expressing neurons to regulate energy balance (38).

SIM1 patients share many features of melanocortin-deficient patients, but notably do not share the accelerated linear growth and increased final height seen in MC4R deficiency (23). Further experiments are needed to elucidate the precise nature of the interaction between MC4R and SIM1 and the common pathways activated by MC4R and SIM1.

The transcriptional targets of SIM1 are unknown; however, the anorexigenic neuropeptide oxytocin is a potential candidate, as the hyperphagia of *Sim1* haploinsufficient mice is partly ameliorated by the central administration of oxytocin and exacerbated by the administration of an oxytocin receptor antagonist (39). PVN oxytocin neurons are activated by the central administration of a selective Mc4r agonist in mice; however, it is unclear whether oxytocin levels are modulated independently, or as a consequence, of defective melanocortin signaling. Impaired oxytocinergic signaling is one mechanism implicated in the hyperphagia and obesity seen in PWS, a clinical syndrome caused by lack of expression of a cluster of maternally imprinted snoRNAs on chromosome 15, which are thought to be involved in alternative mRNA splicing (40). Patients with PWS have a distinctive appearance, with obesity, short stature, small hands, and characteristic facial features. They have developmental delay with prominent behavioral abnormalities, such as obsessive compulsive–type behaviors and emotional lability (40). Some of these features are reminiscent of the cognitive and behavioral deficits seen in SIM1 variant carriers that were varied in manifestation and severity and could be influenced by impaired oxytocinergic signaling. The neurobehavioral features of SIM1 variant carriers are not recognized features of MC4R deficiency, implying that MC4R, SIM1, and OXT are unlikely to interact in an exclusively linear pathway. Further characterization of these features will have implications for the identification of patients in whom variants in SIM1 should be considered and for genetic counseling of SIM1 variant carriers.

Our data demonstrate that rare variants in SIM1 are associated with severe obesity and should be considered in patients with hyperphagic obesity of early onset. The study of patients with SIM1 variants that exhibit reduced activity in vitro provides strong evidence that this transcription factor plays a physiological role in central melanocortin signaling in a specific pathway that regulates food intake independently of energy expenditure, although the precise details of this circuitry remain to be established. Identifying the direct transcriptional targets of SIM1 will be critical to understanding the molecular mechanisms underlying its role in the regulation of food intake and body weight.

Methods

Subjects. The GOOS is a cohort of individuals with severe early onset obesity. Age of obesity onset is less than 10 years. Severe obesity is defined as BMI (weight in kilograms divided by the square of the height in meters) SD score greater than 3. SD scores were calculated according to the UK reference population. 2,100 patients from the cohort of European descent were randomly selected for SIM1 screening. Variants in genes that are known to

cause monogenic obesity were excluded. The Ely study is a prospective population-based cohort study of the etiology and pathogenesis of type 2 diabetes and associated conditions in UK residents of mixed European descent.

Gene screening. The coding region of the human *SIM1* gene (NM 005068) was screened in genomic DNA isolated from whole-blood lymphocytes by direct nucleotide sequencing at the Wellcome Trust Sanger Institute.

Cell culture and generation of stable cell lines. Cell lines were maintained in DMEM (Invitrogen) supplemented with 10% FBS (Gibco; Invitrogen), 10,000 U/ml penicillin, and 10 mg/ml streptomycin (Invitrogen). Doxycycline-inducible stable cell lines were established using the human embryonic kidney 293 Flp-In T-Rex cell system (Invitrogen) as per the manufacturer's instructions.

Plasmid construction. Full-length hSIM1 cDNA, a gift from S.E. Antonarakis (Department of Genetics, University of Geneva, Geneva, Switzerland), was modified to encode a 2xMyc tag at the 3′ end and subcloned into pcDNA5-FRT-TO (Invitrogen). pcDNA5-FRT-TO-hSIM1-Myc was used to generate stable cell lines. Point variants within the hSIM1 coding sequence were introduced into pcDNA5-FRT-TO-hSIM1-Myc using overlap extension PCR. Full details of primer sequences used are included in Supplemental Table 2.

pEF-hARNT-IRES-neo has been previously described (20). pEF-hARNT2-IRES-neo was constructed by amplifying full-length human ARNT2 from 293T cDNA using primers 5′-TAACGCGTCAAGCGGGC-GCCTATC-3′ and 5′-CGTCTAGACTACTCAGAAAACGG-3′, digesting with MluI/XbaI, and ligating into similarly digested pEF-IRES-neo (Genbank Y11035.1). pML/6C-WT was a gift from J. Pelletier (Department of Biochemistry, McGill University, Montreal, Canada).

Transfection and luciferase assays. Cells were transfected with 400 ng pML/6C-WT, 0.5 ng phRL-CMV (Promega), and either 20 ng pEF-hARNT-IRES-neo or 50 ng pEF-hARNT2-IRES-neo using Fugene6 (Roche). Six hours later, cells were induced with doxycycline (Sigma-Aldrich) at 1 μg/ml. Cells were harvested after 16 hours doxycycline treatment and assayed for luciferase activity using the Dual Luciferase Assay System (Promega). Cell lines were assayed in triplicate in at least 3 independent experiments. Firefly luciferase activity was normalized to Renilla luciferase activity for each well (relative luciferase units), and the 3 values were averaged. In the case of the 2 WT lines, values for all 6 wells were averaged. These data were converted to log values for the purposes of statistical analysis.

Immunoblotting. Following 16-hour doxycycline treatment (1 μg/ml), whole-cell extracts were prepared as previously described (41). Samples were subjected to SDS-PAGE, transferred to nitrocellulose, and immunoblotted using anti-Myc 4A6 antibody (Millipore) and anti-ARNT2 M-165 l antibody (Santa Cruz Biotechnology Inc.). Anti-α-tubulin antibody MCA78G (AbD Serotec) was used to detect tubulin as a loading control.

Immunocytochemistry. Cells were grown on glass coverslips coated with gelatin and treated with 1 μg/ml doxycycline for 16 hours prior to fixing with 4% PFA. Fixed cells were stained with α-Myc-tag primary antibody (4A6, 1:500; Millipore) and Alexa Fluor 594 α-mouse secondary antibody (1:1000). Coverslips were mounted with ProLong Gold antifade reagent with DAPI (Invitrogen), and images were taken with a Zeiss Axioplan2 microscope.

Homology modeling. Homology modeling was carried out in the ICM-Pro program suite using the homology add-on (42). Each model of the heterodimer was first created individually using the sequence for human SIM1 residues 1–225 (Uniprot number P81133) and for human ARNT2 residues 1–330 (Uniprot number Q9HBZ2) to include the bHLH and PASA domains of each protein. Homologous structures were searched for in the pdb, and the CLOCK-BMAL structure was the closest homologous structure (PDB:4F3L) (21). A homology model of SIM1 was created using the CLOCK protein as a 3D template, and the ARNT2 model was created using BMAL as a 3D model. After creation of the individual models, they were subject to regularization and model refinement within ICM-Pro (to optimize model geometry, carry out energy minimization, and alleviate clashing side-chains).

The 2 structures were then docked to using the CLOCK-BMAL structure as a guide. Further regularization and model refinement was carried out to ensure the integrity of the dimer interface. Finally, several loops were subject to loop modeling to improve clashing at the dimer interface using the ICM-Pro loop modeling utility (43). Figure 3, D–F was created using PyMOL (The PyMOL Molecular Graphics System, Version 1.2r3pre; Schrödinger, LLC).

Metabolic phenotyping of patients. Subjects were invited to participate in clinical studies at the Wellcome Trust Clinical Research Facility at Addenbrooke's Hospital (Cambridge, UK). Weight and height were measured barefoot in light clothing. DEXA (DPX software; Lunar Corp) was used to determine body composition. Ad libitum energy intake was assessed using a breakfast meal of known macronutrient content after an overnight fast and was expressed per kilogram of lean body mass as measured by DEXA to allow comparison between individuals of different body weights and compositions. BMR and respiratory quotient were determined by indirect calorimetry after an overnight fast using an open circuit, ventilated canopy measurement system (Europa Gas Exchange Monitor; NutrEn Technology Ltd.). After adjustment for body composition, BMR was compared with predicted metabolic rate based on age- and sex-specific equations. Blood pressure was measured using automated brachial (DINAMAP; GE Healthcare) or wrist (OMRON Healthcare) monitors. Heart rate was recorded using a wearable sensor (Actiheart; CamNtech Ltd.) from which heart rate variability parameters were derived from cleaned inter-beat interval time series (44).

Statistics. For reporter assays, univariate ANOVA was performed on the log values generated for each mutant with reference to WT using SPSS 17.0. For clinical measurements, values were compared using the unpaired 2-tailed Student's t test ($P < 0.05$ deemed significant).

Study approval. All studies were approved by the Cambridge Regional Ethics Committee and were conducted in accordance with the principles of the Declaration of Helsinki. Each subject, or his or her parent for those under 16, provided written informed consent; minors provided oral consent.

Address correspondence to: I. Sadaf Farooqi, University of Cambridge Metabolic Research Laboratories and NIHR Cambridge Biomedical Research Centre, Institute of Metabolic Science, Box 289, Addenbrooke's Hospital, Hills Road, Cambridge, CB2 0QQ, United Kingdom. Phone: 44.1223.762634; Fax: 44.1223.762657; E-mail: isf20@cam.ac.uk.

1. Farooqi IS, O'Rahilly S. Genetic factors in human obesity. *Obes Rev*. 2007;1(1):37–40.
2. Michaud JL, Rosenquist T, May NR, Fan CM. Development of neuroendocrine lineages requires the bHLH-PAS transcription factor SIM1. *Genes Dev*. 1998;12(20):3264–3275.
3. Michaud JL, et al. Sim1 haploinsufficiency causes hyperphagia, obesity and reduction of the paraventricular nucleus of the hypothalamus. *Hum Mol Genet*. 2001;10(14):1465–1473.
4. Kublaoui BM, Holder JL Jr, Gemelli T, Zinn AR. Sim1 haploinsufficiency impairs melanocortin-mediated anorexia and activation of paraventricular nucleus neurons. *Mol Endocrinol*. 2006; 20(10):2483–2492.
5. Holder JL. et al. Sim1 gene dosage modulates the homeostatic feeding response to increased dietary fat in mice. *Am J Physiol Endocrinol Metab*. 2004; 287(1):E105–E113.
6. Tolson KP, Gemelli T, Gautron L, Elmquist JK, Zinn AR, Kublaoui BM. Postnatal Sim1 deficiency causes hyperphagic obesity and reduced Mc4r and oxytocin expression. *J Neurosci*. 2010;30(10):3803–3812.
7. Kublaoui BM, Holder JL Jr, Tolson KP, Gemelli T, Zinn AR. SIM1 overexpression partially rescues agouti yellow and diet-induced obesity by normalizing food intake. *Endocrinology*. 2006;147(10):4542–4549.
8. Traurig M, et al. Common variation in SIM1 is reproducibly associated with BMI in Pima Indians. *Diabetes*. 2009;58(7):1682–1689.
9. Ahituv N, et al. Medical sequencing at the extremes of human body mass. *Am J Hum Genet*. 2007; 80(4):779–791.
10. Faivre L, et al. Deletion of the SIM1 gene (6q16.2) in a patient with a Prader-Willi-like phenotype. *J Med Genet*. 2002;39(8):594–596.
11. Holder JL Jr, Butte NF, Zinn AR. Profound obesity associated with a balanced translocation that disrupts the SIM1 gene. *Hum Mol Genet*. 2000; 9(1):101–108.
12. Hung CC, et al. Studies of the SIM1 gene in relation to human obesity and obesity-related traits. *Int J Obes (Lond)*. 2007;31(3):429–434.
13. Chapman-Smith A, Lutwyche JK, Whitelaw ML. Contribution of the Per/Arnt/Sim (PAS) domains to DNA binding by the basic helix-loop-helix PAS transcriptional regulators. *J Biol Chem*. 2004; 279(7):5353–5362.
14. Crews ST, Thomas JB, Goodman CS. The Drosophila single-minded gene encodes a nuclear protein with sequence similarity to the per gene product. *Cell*. 1988;52(1):143–151.
15. Moffett P, Pelletier J. Different transcriptional properties of mSim-1 and mSim-2. *FEBS Lett*. 2000; 466(1):80–86.
16. Michaud JL, DeRossi C, May NR, Holdener BC, Fan CM. ARNT2 acts as the dimerization partner of SIM1 for the development of the hypothalamus. *Mech Dev*. 2000;90(2):253–261.
17. Wharton KA Jr, Franks RG, Kasai Y, Crews ST. Control of CNS midline transcription by asymmetric E-box-like elements: similarity to xenobiotic responsive regulation. *Development*. 1994; 120(12):3563–3569.
18. Metz RP, Kwak HI, Gustafson T, Laffin B, Porter WW. Differential transcriptional regulation by mouse single-minded 2s. *J Biol Chem*. 2006; 281(16):10839–10848.
19. Yamaki A, Kudoh J, Shimizu N, Shimizu Y. A novel nuclear localization signal in the human single-minded proteins SIM1 and SIM2. *Biochem Biophys Res Commun*. 2004;313(3):482–488.
20. Scheuermann TH, Tomchick DR, Machius M, Guo Y, Bruick RK, Gardner KH. Artificial ligand binding within the HIF2alpha PAS-B domain of the HIF2 transcription factor. *Proc Natl Acad Sci U S A*. 2009;106(2):450–455.
21. Huang N, et al. Crystal structure of the heterodimeric CLOCK:BMAL1 transcriptional activator complex. *Science*. 2012;337(6091):189–194.
22. King HA, Hoelz A, Crane BR, Young MW. Structure of an enclosed dimer formed by the Drosophila period protein. *J Mol Biol*. 2011;413(3):561–572.
23. Martinelli CE, et al. Obesity due to melanocortin 4 receptor (MC4R) deficiency is associated with increased linear growth and final height, fasting hyperinsulinemia, and incompletely suppressed growth hormone secretion. *J Clin Endocrinol Metab*. 2011;96(1):E181–E188.
24. Greenfield JR, et al. Modulation of blood pressure by central melanocortinergic pathways. *N Engl J Med*. 2009;360(1):44–52.
25. Ghoussaini M, et al. Analysis of the SIM1 contribution to polygenic obesity in the French population. *Obesity (Silver Spring)*. 2010;18(8):1670–1675.
26. Partch CL, Gardner KH. Coactivators necessary for transcriptional output of the hypoxia inducible factor, HIF, are directly recruited by ARNT PAS-B. *Proc Natl Acad Sci U S A*. 2011;108(19):7739–7744.
27. Sato TK, et al. Feedback repression is required for mammalian circadian clock function. *Nat Genet*. 2006;38(3):312–319.
28. Woods SL, Whitelaw ML. Differential activities of murine single minded 1 (SIM1) and SIM2 on a hypoxic response element. Cross-talk between basic helix-loop-helix/per-Arnt-Sim homology transcription factors. *J Biol Chem*. 2002;277(12):10236–10243.
29. Furness SG, Lees MJ, Whitelaw ML. The dioxin (aryl hydrocarbon) receptor as a model for adaptive responses of bHLH/PAS transcription factors. *FEBS Lett*. 2007;581(19):3616–3625.
30. Farrall AL, Whitelaw ML. The HIF1alpha-inducible pro-cell death gene BNIP3 is a novel target of SIM2s repression through cross-talk on the hypoxia response element. *Oncogene*. 2009;28(41):3671–3680.
31. Woods S, Farrall A, Procko C, Whitelaw ML. The bHLH/Per-Arnt-Sim transcription factor SIM2 regulates muscle transcript myomesin2 via a novel, non-canonical E-box sequence. *Nucleic Acids Res*. 2008;36(11):3716–3727.
32. Ema M, et al. Two new members of the murine Sim gene family are transcriptional repressors and show different expression patterns during mouse embryogenesis. *Mol Cell Biol*. 1996;16(10):5865–5875.
33. Bonnefond A, et al. Loss-of-function mutations in *SIM1* contribute to obesity and Prader-Willi-like features. *J Clin Invest*. 2013;123(7):3037–3041.
34. Xu C, Fan CM. Allocation of paraventricular and supraoptic neurons requires Sim1 function: a role for a Sim1 downstream gene PlexinC1. *Mol Endocrinol*. 2007;21(5):1234–1245.
35. Duplan SM, Boucher F, Alexandrov L, Michaud JL. Impact of Sim1 gene dosage on the development of the paraventricular and supraoptic nuclei of the hypothalamus. *Eur J Neurosci*. 2009; 30(12):2239–2249.
36. Yang C, et al. Adenoviral-mediated modulation of Sim1 expression in the paraventricular nucleus affects food intake. *J Neurosci*. 2006;26(26):7116–7120.
37. Balthasar N, et al. Divergence of melanocortin pathways in the control of food intake and energy expenditure. *Cell*. 2005;123(3):493–505.
38. Xi D, Gandhi N, Lai M, Kublaoui BM. Ablation of Sim1 neurons causes obesity through hyperphagia and reduced energy expenditure. *PLoS One*. 2012;7(4):e36453.
39. Kublaoui BM, Gemelli T, Tolson KP, Wang Y, Zinn AR. Oxytocin deficiency mediates hyperphagic obesity of Sim1 haploinsufficient mice. *Mol Endocrinol*. 2008;22(7):1723–1734.
40. Goldstone AP. Prader-Willi syndrome: advances in genetics, pathophysiology and treatment. *Trends Endocrinol Metab*. 2004;15(1):12–20.
41. Whitelaw M, Pongratz I, Wilhelmsson A, Gustafsson JA, Poellinger L. Ligand-dependent recruitment of the Arnt coregulator determines DNA recognition by the dioxin receptor. *Mol Cell Biol*. 1993; 13(4):2504–2514.
42. Cardozo T, Totrov M, Abagyan R. Homology modeling by the ICM method. *Proteins*. 1995;23(3):403–414.
43. Arnautova YA, Abagyan RA, Totrov M. Development of a new physics-based internal coordinate mechanics force field and its application to protein loop modeling. *Proteins*. 2011;79(2):477–498.
44. Brage S, Brage N, Franks PW, Ekelund U, Wareham NJ. Reliability and validity of the combined heart rate and movement sensor Actiheart. *Eur J Clin Nutr*. 2005;59(4):561–570.