

Mutations in *SGOL1* cause a novel cohesinopathy affecting heart and gut rhythm

Philippe Chetaille^{1,14}, Christoph Preuss^{2,14}, Silja Burkhard^{3,14}, Jean-Marc Côté¹, Christine Houde¹, Julie Castilloux¹, Jessica Piché², Natacha Gosset², Séverine Leclerc², Florian Wünnemann², Maryse Thibeault², Carmen Gagnon², Antonella Galli⁴, Elizabeth Tuck⁴, Gilles R Hickson⁵, Nour El Amine⁵, Ines Boufaied⁵, Emmanuelle Lemyre⁵, Pascal de Santa Barbara⁶, Sandrine Faure⁶, Anders Jonzon⁷, Michel Cameron², Harry C Dietz⁸, Elena Gallo-McFarlane⁸, D Woodrow Benson⁹, Claudia Moreau⁵, Damian Labuda⁵, FORGE Canada Consortium¹⁰, Shing H Zhan¹¹, Yaoqing Shen¹¹, Michèle Jomphe¹², Steven J M Jones¹¹, Jeroen Bakkers³ & Gregor Andelfinger^{2,5,13}

The pacemaking activity of specialized tissues in the heart and gut results in lifelong rhythmic contractions. Here we describe a new syndrome characterized by Chronic Atrial and Intestinal Dysrhythmia, termed CAID syndrome, in 16 French Canadians and 1 Swede. We show that a single shared homozygous founder mutation in *SGOL1*, a component of the cohesin complex, causes CAID syndrome. Cultured dermal fibroblasts from affected individuals showed accelerated cell cycle progression, a higher rate of senescence and enhanced activation of TGF- β signaling. Karyotypes showed the typical railroad appearance of a centromeric cohesion defect. Tissues derived from affected individuals displayed pathological changes in both the enteric nervous system and smooth muscle. Morpholino-induced knockdown of *sgo1* in zebrafish recapitulated the abnormalities seen in humans with CAID syndrome. Our findings identify CAID syndrome as a novel generalized dysrhythmia, suggesting a new role for *SGOL1* and the cohesin complex in mediating the integrity of human cardiac and gut rhythm.

Disturbances of pacemaker activity in the heart and gut can have varied clinical manifestations. Dysregulation of the cardiac sinus node results in sick sinus syndrome (SSS), the most common cause of pacemaker implantation¹. SSS is rare in children and is characterized by persistently decreased heart rhythm, episodes of sinoatrial block and/or chronotropic incompetence. In the gut, pacemaking is mediated through the network of interstitial cells of Cajal and the autonomous enteric nervous system (ENS). Chronic intestinal pseudo-obstruction

(CIPO) is a rare and severe disorder of gastrointestinal motility in which intestinal obstruction occurs in the absence of a mechanical obstacle. Several studies identified genetic risk factors for SSS^{2–8} and CIPO^{9,10}, but in most cases the genetic causes remain enigmatic.

We identified 17 subjects in whom SSS and CIPO co-occurred during the first four decades of life (Fig. 1a). Distinctive clinical features included atrial dysrhythmias, SSS and valve anomalies, as well as a requirement for pacemaker implantation and total parenteral nutrition (Fig. 1b,c and Supplementary Tables 1 and 2). In different individuals with CAID, CIPO was found to be of neurogenic or myogenic origin (Supplementary Fig. 1). No case had clinical evidence of other congenital anomalies or manifestations of known cohesinopathies. Family evaluation suggested inheritance in an autosomal recessive manner, prompting us to perform whole-exome sequencing of three unrelated probands. We identified only one homozygous pathogenic variant shared by all probands, namely, c.67A>G, n.367–2014T>C, in *SGOL1* (rs199815268), which was predicted to encode a damaging change, p.Lys23Glu, at a highly conserved residue (Supplementary Fig. 2a–c and Supplementary Table 3). The mutation was extremely rare (minor allele frequency (MAF) < 1%) in public databases¹¹ (see URLs) and was absent in 360 French-Canadian control exomes. We confirmed the homozygous mutation by Sanger sequencing in all 14 surviving French-Canadian cases and the 1 Swedish case (Fig. 1a). The mutation was absent from 11 isolated pediatric SSS cases without CIPO and 43 isolated pediatric CIPO cases without SSS.

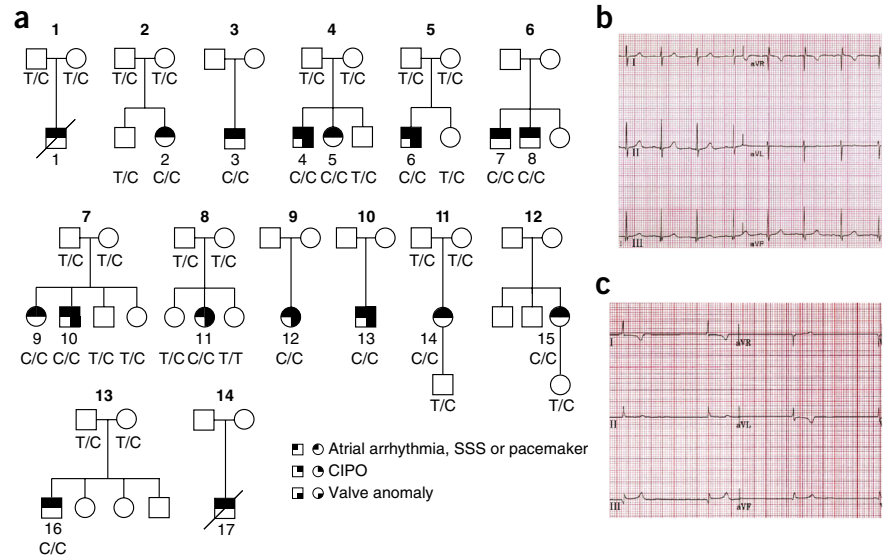
To genetically fine map the disease-causing haplotype and exclude the presence of copy number variants (CNVs), we performed Illumina

¹Department of Pediatrics, Centre Mère Enfants Soleil, Centre Hospitalier de l'Université (CHU) de Québec, Quebec City, Quebec, Canada. ²Cardiovascular Genetics, Department of Pediatrics, Centre Hospitalier Universitaire Sainte-Justine Research Centre, Université de Montréal, Montreal, Quebec, Canada.

³Hubrecht Institute, Royal Netherlands Academy of Arts and Sciences (KNAW) and University Medical Center Utrecht, Utrecht, the Netherlands. ⁴Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK. ⁵Department of Pediatrics, Université de Montréal, Montreal, Quebec, Canada. ⁶INSERM U1046, Montpellier, France. ⁷Department of Women's and Children's Health, Section for Pediatrics, Astrid Lindgren's Children's Hospital, Uppsala University, Uppsala, Sweden. ⁸Howard Hughes Medical Institute, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. ⁹Department of Pediatrics, Medical College of Wisconsin, Milwaukee, Wisconsin, USA. ¹⁰Membership of the Steering Committee for the Consortium is given in the Acknowledgments. ¹¹Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia, Canada. ¹²Projet BALSAC, Université du Québec à Chicoutimi, Chicoutimi, Quebec, Canada. ¹³Department of Biochemistry, Université de Montréal, Montreal, Quebec, Canada. ¹⁴These authors contributed equally to this work. Correspondence should be addressed to G.A. (gregor.andelfinger@recherche-ste-justine.qc.ca).

Received 8 April; accepted 11 September; published online 5 October 2014; doi:10.1038/ng.3113

Figure 1 Synopsis of families and electrocardiographic features. (a) Pedigrees of the CAID syndrome cohort. Wild-type (T/T), heterozygous (T/C) and homozygous (C/C) mutation status are shown. Family and subject numbers refer to those used in the text and **Supplementary Table 2**. In family 1, in which the index case was deceased without any available biological material, we found both parents to be heterozygous for the CAID-linked mutation in *SGOL1*. (b,c) Evolution of clinical electrocardiographic findings in a CAID case, showing normal sinus rhythm at 3 years of age (b) and sinus node dysfunction with junctional escape rhythm (40 beats per minute) at 19 years of age (c). I, II, III, aVR, aVL and aVF denote corresponding peripheral leads.



HumanOmni5-Quad genotyping for 13 cases and 3 unaffected family members. Homozygosity mapping with PLINK and HomozygosityMapper^{12,13} (see URLs) identified a single 1.4-Mb region between SNPs rs2929378 and rs442920 harboring *SGOL1* that was identical in all the cases (**Supplementary Fig. 3**). Further haplotype analysis identified a rare 700-kb disease-associated haplotype shared by the Swedish and French-Canadian cases lacking rare CNVs. Genealogy and genetic analyses excluded any direct relatedness for the Swedish and French-Canadian individuals (PLINK pi-hat < 0.2; *F* inbreeding coefficient -0.1 -0). Our population genetic analysis determined that the rare disease-causing haplotype was of northern European origin and demonstrated the absence of tagging SNPs for this haplotype in Asian, African and admixed

American populations (**Supplementary Fig. 4**). Reconstitution of ascending genealogies from the BALSAC population database¹⁴ for 8 French-Canadian cases identified 64 common ancestors married in the seventeenth century in France or in Quebec City. Identity-by-descent analysis supported the observation that all the French-Canadian CAID cases derived from the eastern part of Quebec and shared a common ancestral haplotype for *SGOL1*, whose age was estimated at 13 ± 4 generations. Allele dropping analysis identified a founder couple, married in France in 1620, whose likelihood of being the mutation carriers exceeded that for other founders by a factor

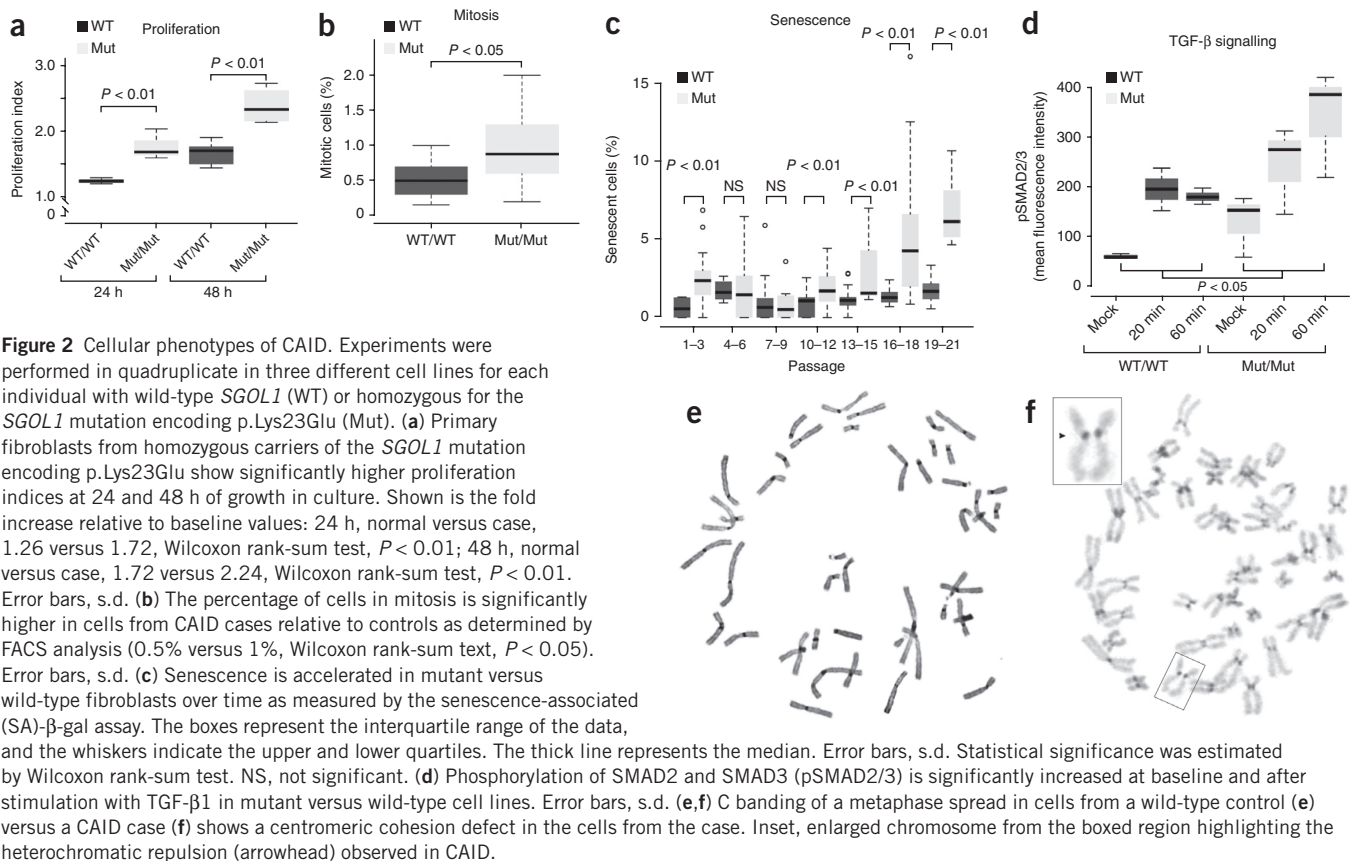
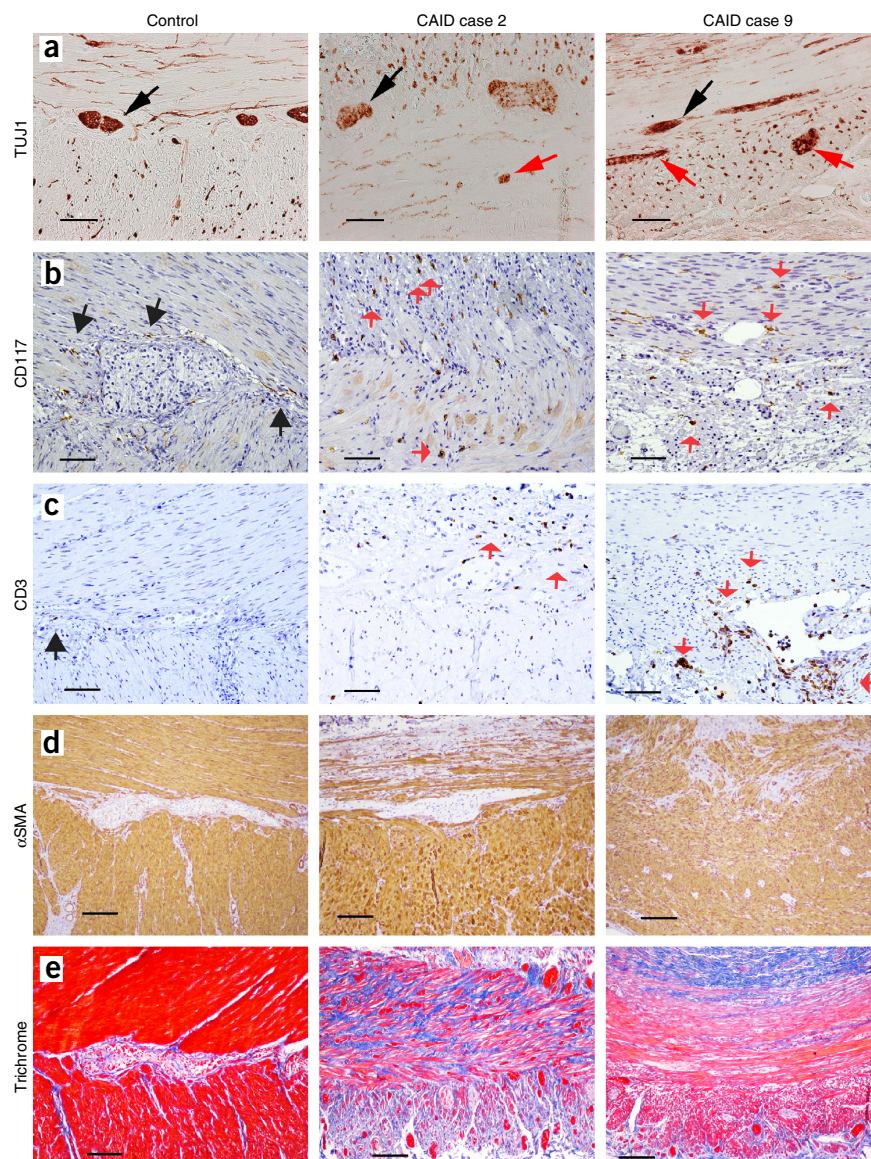


Figure 2 Cellular phenotypes of CAID. Experiments were performed in quadruplicate in three different cell lines for each individual with wild-type *SGOL1* (WT) or homozygous for the *SGOL1* mutation encoding p.Lys23Glu (Mut). (a) Primary fibroblasts from homozygous carriers of the *SGOL1* mutation encoding p.Lys23Glu show significantly higher proliferation indices at 24 and 48 h of growth in culture. Shown is the fold increase relative to baseline values: 24 h, normal versus case, 1.26 versus 1.72, Wilcoxon rank-sum test, $P < 0.01$; 48 h, normal versus case, 1.72 versus 2.24, Wilcoxon rank-sum test, $P < 0.01$. Error bars, s.d. (b) The percentage of cells in mitosis is significantly higher in cells from CAID cases relative to controls as determined by FACS analysis (0.5% versus 1%, Wilcoxon rank-sum test, $P < 0.05$). Error bars, s.d. (c) Senescence is accelerated in mutant versus wild-type fibroblasts over time as measured by the senescence-associated (SA)-β-gal assay. The boxes represent the interquartile range of the data, and the whiskers indicate the upper and lower quartiles. The thick line represents the median. Error bars, s.d. Statistical significance was estimated by Wilcoxon rank-sum test. NS, not significant. (d) Phosphorylation of SMAD2 and SMAD3 (pSMAD2/3) is significantly increased at baseline and after stimulation with TGF-β1 in mutant versus wild-type cell lines. Error bars, s.d. (e,f) C banding of a metaphase spread in cells from a wild-type control (e) versus a CAID case (f) shows a centromeric cohesion defect in the cells from the case. Inset, enlarged chromosome from the boxed region highlighting the heterochromatic repulsion (arrowhead) observed in CAID.

Figure 3 Synopsis of the gastrointestinal histology of CAID. Full-thickness biopsies from a wild-type control and two individuals with CAID. All biopsies are oriented to show the circular muscular layer on top and the longitudinal layer on the bottom. (a–e) Staining was performed for TUJ1 (neuron-specific β III tubulin, a neuronal marker expressed in intestinal ganglial cells) (a), CD117 (c-KIT, a marker of Cajal cells) (b), CD3 (a marker of mature lymphocytes) (c), α SMA (α smooth muscle actin, a smooth muscle cell marker) (d) and Trichrome-Masson (distinguishes cells (red) from surrounding connective tissue (blue)) (e). (a) In addition to the presence of normal intestinal ganglia (black arrows), smaller ganglia mislocalized into the longitudinal smooth muscle layer are seen in CAID cases (red arrows). (b) Whereas Cajal cells are localized in a periganglial fashion in wild-type cells (black arrows), they are seen abundantly within the circular and longitudinal smooth muscle layers in CAID (red arrows). (c) Mature lymphocytes abundantly invade the smooth muscle layers in CAID (red arrows) in comparison to wild-type tissue (black arrow). (d) α SMA staining demonstrates the presence of the two smooth muscle layers but also indicates vacuolization and disruption of the fiber architecture of smooth muscles previously reported in individuals with CIPO. (e) This disruption of wall architecture is further confirmed by Trichrome-Masson staining, which also documents extensive fibrotic remodeling of the intestinal wall. Scale bars, 100 μ m (a–c) and 200 μ m (d,e).



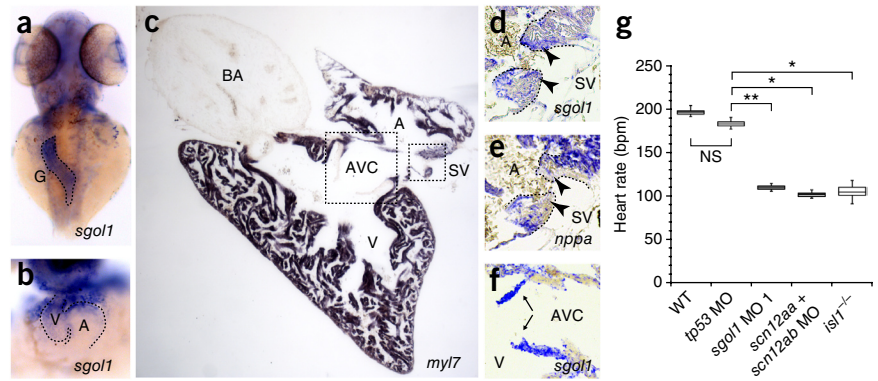
of >100 (Supplementary Fig. 5a,b). The haplotype sharing between the Swedish and French-Canadian cases (Supplementary Fig. 5c) suggests that they have common ancestry at about 30 generations, or 900 years, ago¹⁵. Taken together, our analyses support the association of *SGOL1* with the disease and suggest a transatlantic founder effect in the seventeenth century when Nouvelle France was settled.

Because of the known involvement of *SGOL1* in cell division, we studied cell cycling parameters in primary fibroblast cultures from three homozygous carriers of the *SGOL1* mutation encoding p.Lys23Glu. These fibroblasts exhibited significantly faster cell proliferation ($P < 0.01$) due to a higher percentage of cells in M phase ($P < 0.05$), as determined by intracellular flow cytometry, than corresponding cells from healthy controls (Fig. 2a,b). Over time, the rate of senescence was higher in mutant versus control cell lines (Fig. 2c). In control fibroblast cell lines, the *SGOL1* protein was restricted almost exclusively to the nucleus, where it formed a distinct punctate pattern. In contrast, the mutated *SGOL1* protein showed a homogenous distribution inside the nucleus of case-derived fibroblasts as well as the cytoplasm (Supplementary Fig. 6). This altered organization of nuclear *SGOL1* was also apparent in quantitative analysis of the signal intensities for nuclei (Supplementary Fig. 7). During mitosis, the mutant form of *SGOL1* was localized in an ordered fashion around the centromeres but displayed a rather homogenous cytoplasmic localization pattern (Supplementary Fig. 6). Because transforming growth factor (TGF)- β signaling is a known mitogenic cascade with

clinical relevance to arrhythmias, we assayed the behavior of fibroblasts in response to TGF- β stimulation¹⁶. In comparison to wild-type fibroblast cell cultures, cell lines homozygously expressing mutant *SGOL1* had higher baseline levels of phosphorylation of SMAD2 and SMAD3 (SMAD2/3), which constitute a critical downstream mediator of the TGF- β signaling pathway. These fibroblasts also showed higher levels of phosphorylated SMAD2/3 in response to TGF- β 1 stimulation (Fig. 2d). On the basis of the known function of *SGOL1*, we assumed that there might be a potentially informative karyotype in the metaphase spreads of chromosomes from cases. Indeed, C banding identified heterochromatin repulsion at the centromere (resulting in a railroad track appearance) without signs of aneuploidy, fostering the diagnosis of a novel cohesinopathy (Fig. 2e,f).

In normal tissues, *SGOL1* is ubiquitously expressed in the intestinal wall, including in smooth muscle and the ENS (Supplementary Fig. 8). CAID cases showed characteristic anomalies in the ENS, which controls gastrointestinal reflexes. We found both hypoplastic ganglia and mislocalization of ganglia and Cajal cells in the circular and longitudinal smooth muscle layers, reminiscent of the phenotypes described in mouse models with altered *Shh* gradients^{17,18} (Fig. 3a,b). Compatible with the increased TGF- β signaling seen in cell culture,

Figure 4 Spatiotemporal expression and function of *sgol1* in zebrafish. (a,b) *In situ* hybridization for *sgol1* on wild-type zebrafish embryos (3 days post-fertilization (d.p.f.)). (a) Dorsal view; *sgol1* expression marks the intestinal bulb and intestine (dotted line). (b) Ventral view; *sgol1* expression in the embryonic heart (dotted line). (c) Labeling with the myocardial marker *myl7* delineates the myocardium; boxed areas indicate the magnified regions shown in d–f. (d–f) Section *in situ* hybridization on adult wild-type zebrafish hearts. (d,e) Expression of *sgol1* at the base of the venous valves (demarcated area) in the sinoatrial region (d). Expression of *nppa* in the sinoatrial region (e). Arrowheads indicate *nppa*-negative cells located at the base of the venous valves. (f) *sgol1* expression in the atrioventricular valves (arrows). (g) Analysis of heart rate by high-speed video imaging. Box-whisker plots show the cardiac cycle length with the indicated morpholino injections ($n = 15$ embryos per group). Significance was determined by Student's *t* test: * $P < 0.05$, ** $P < 0.01$, NS $P > 0.05$. A, atrium; AVC, atrioventricular canal; BA, bulbus arteriosus; G, gut; SV, sinus venosus; V, ventricle; bpm, beats per minute.



we found abundant T cells in the intestinal smooth muscle layers (Fig. 3c). Furthermore, thinning of smooth muscle layers, fragmentation of the smooth muscle fiber architecture and extensive fibrosis, a hallmark of chronic TGF- β activation, were observed (Fig. 3d,e). Together, these data suggest that both neuromuscular alterations contribute to the CAID syndrome.

To illustrate the role of *sgol1* during heart and gut development, we established a zebrafish model. We observed a distinct spatiotemporal expression pattern for zebrafish *sgol1* in developing heart and gut (Fig. 4a–d,f). Notably, a strong and discrete expression signal was found at the level of the sinus venosus from which the pacemaker tissue arises, in a mutually exclusive pattern with *nppa*, a marker of terminal cardiomyocyte differentiation (Fig. 4c–f)¹⁹. Morpholino-mediated knockdown of *sgol1* significantly decreased heart rate in zebrafish ($P < 0.0001$), with an effect size similar to the knockdown of known conduction and pacemaker regulators such as *scn12a* (the zebrafish ortholog of *SCN5A*) and *isl1* (Fig. 4g and Supplementary Fig. 9)^{2,20}.

All the affected individuals described in this study shared a unique combination of cardiac arrhythmias and intestinal pseudo-obstruction that defines a novel cohesinopathy syndrome we have termed CAID. Several lines of genetic and functional evidence show that a single homozygous mutation in *SGOL1* predisposes to this generalized co-occurring cardiac and intestinal rhythm phenotype. This mutation arose recently on a rare haplotype in the European population, with an estimated age of 900 years. The identification of a Swedish case sharing an extended haplotype with the French-Canadian cases highlights the possibility that migration might have had an important role in disease etiology. Linking the geographical gradient of the *SGOL1* haplotype (Supplementary Fig. 4) and the known migration routes of Vikings in the twelfth century from Scandinavia to Normandy supports the idea of a transatlantic founder effect originating in Scandinavia.

We propose that the mutated *SGOL1* protein results in loss of cohesin protection, accelerating cell cycling and senescence and thereby replicating the bradycardia-tachycardia syndrome usually observed in elderly individuals. Our novel syndrome falls under the umbrella term cohesinopathy, which includes phenotypically diverse entities such as premature ovarian failure (*STAG3*), Cornelia de Lange syndrome (*NIPBL*, *RAD21*, *SMC1* and *HDAC8*), Warsaw breakage syndrome (*DDX11*), Nijmegen breakage syndrome (*NBS1*) and Roberts syndrome (*ESCO2*), which all result from mutations in cohesin components^{21–27}. The unique CAID phenotype raises the possibility that both loss-of-function and gain-of-function effects

co-occur. The disruption of the known interaction between *SGOL1* and PP2A²⁸ might result in loss of separase inhibition at the nuclear level, whereas mislocalization of *SGOL1* in the cytosol during interphase might alter PP2A-mediated ion channel regulation²⁹. We hypothesize that spatiotemporally restricted expression of *SGOL1* during embryogenesis or the restricted requirement for *SGOL1* in adult pacemakers might confer the target organ specificity observed in CAID syndrome.

The involvement of *SGOL1* in the cohesin complex could entail consequences for long-range transcriptional regulation^{24,30}, possibly interfering with the expression of gene products associated with SSS or CIPO^{2,3,6,8,9}. Notably, the histological findings in individuals with CAID might point to an interaction of both developmental (mislocalized, abnormal ganglia) and acquired (destruction of smooth muscle architecture, fibrosis) phenotypes. Thus, in addition to acting through TGF- β signaling, a non-canonical role for *SGOL1* and the cohesin complex in maintaining shared transcriptional modules of the sinoatrial node in the heart and Cajal cells in the gut is an intriguing possibility. In conclusion, the findings in this study support a hitherto unsuspected role for *SGOL1* in the lifelong integrity of human pacemaker activity.

URLs. HomozygosityMapper, <http://www.homozygositymapper.org/>; PLINK, <http://pngu.mgh.harvard.edu/~purcell/plink/>; BALSAC, <http://balsac.uqac.ca/>; Exome Variant Server (EVS), <http://evs.gs.washington.edu/EVS/>; SIFT, <http://sift.jcvi.org/>; SeattleSeq, <http://snpgs.washington.edu/>; R GenLib software package, <http://balsac.uqac.ca/english/acces-aux-donnees-2/service-aux-chercheurs/>.

METHODS

Methods and any associated references are available in the [online version of the paper](#).

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We are deeply indebted to all participating families and patients. We thank L. Gosselin, L. L'Écuyer, M.-F. Boudreault, C. Faure, F. LeDeist, L. D'Aoust and the McGill Genome Québec Innovation Centre (D. Vincent) for expert assistance. We gratefully acknowledge the contribution of all referring physicians and supporting staff to this project. We would like to thank J. Marcadier (clinical coordinator) and C. Beaulieu (project manager) for their contributions to the infrastructure of the FORGE Canada Consortium. The FORGE (Finding of Rare Disease Genes)

Consortium Steering Committee includes Kym Boycott (University of Ottawa), Jan Friedman (University of British Columbia), Jacques Michaud (Université de Montréal), François Bernier (University of Calgary), Michael Brudno (University of Toronto), Bridget Fernandez (Memorial University), Bartha Knoppers (McGill University) and Steve Scherer (University of Toronto). FORGE Canada was funded by the Canadian government through Genome Canada, the Canadian Institutes of Health Research and the Ontario Genomics Institute (OGI-049). Additional funding was provided by Génome Québec, Genome British Columbia, Fondation Nussia and André Aisenstadt, La Fondation du Grand défi Pierre Lavoie, Fondation CHU Sainte-Justine, Fondation Leducq, Association des Pseudo Obstructions Intestinales Chroniques (France), the Netherlands Organization for Scientific Research (NWO) and ZonMw (91212086). G.A. holds a Senior Scholarship from the Fonds de Recherche du Québec-Santé and holds the Banque Nationale Research Chair in Cardiovascular Genetics. We thank G. Rouleau for sharing whole-exome data for control purposes and M. Samuels for critical reading of the manuscript.

AUTHOR CONTRIBUTIONS

P.C., J.-M.C. and G.A. identified the novel CAID syndrome in the families. Functional experiments were performed by F.W., N.G., J.P., S.L., C.G., E.L., I.B., J.B., S.B., A.G., E.T., S.F., M.C., N.E.A., E.G.-M., S.H.Z., Y.S., M.J. and S.J.M.J. Further patient recruitment and enrollment was performed by P.C., J.-M.C., C.H., J.C., A.J., P.d.S.B., M.T., D.W.B. and G.A. C.P. performed and supervised the genotyping and the haplotype and population genetic analysis. G.R.H., G.A., J.P. and N.E.A. performed confocal microscopy for the cell lines and analyzed the data. S.B. and J.B. established the zebrafish model. C.G., J.P. and E.G.-M. performed the functional cell line experiments. Genealogical analysis was provided by M.J., D.L. and C.M. H.C.D. provided advice on the analysis of the functional data derived from the cell lines. P.C., C.P., J.B., S.J.M.J. and G.A. wrote the manuscript with contributions and input from all authors. G.A. designed and supervised the study.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Rodriguez, R.D. & Schocken, D.D. Update on sick sinus syndrome, a cardiac disorder of aging. *Geriatrics* **45**, 26–30, 33–36 (1990).
- Benson, D.W. *et al.* Congenital sick sinus syndrome caused by recessive mutations in the cardiac sodium channel gene (*SCN5A*). *J. Clin. Invest.* **112**, 1019–1028 (2003).
- Laish-Farkash, A. *et al.* A novel mutation in the *HCN4* gene causes symptomatic sinus bradycardia in Moroccan Jews. *J. Cardiovasc. Electrophysiol.* **21**, 1365–1372 (2010).
- Milanesi, R., Baruscotti, M., Gnecci-Ruscone, T. & DiFrancesco, D. Familial sinus bradycardia associated with a mutation in the cardiac pacemaker channel. *N. Engl. J. Med.* **354**, 151–157 (2006).
- Schulze-Bahr, E. *et al.* Pacemaker channel dysfunction in a patient with sinus node disease. *J. Clin. Invest.* **111**, 1537–1545 (2003).
- Holm, H. *et al.* A rare variant in *MYH6* is associated with high risk of sick sinus syndrome. *Nat. Genet.* **43**, 316–320 (2011).
- Mohler, P.J. *et al.* Ankyrin-B mutation causes type 4 long-QT cardiac arrhythmia and sudden cardiac death. *Nature* **421**, 634–639 (2003).
- Mohler, P.J. *et al.* A cardiac arrhythmia syndrome caused by loss of ankyrin-B function. *Proc. Natl. Acad. Sci. USA* **101**, 9137–9142 (2004).
- Gargiulo, A. *et al.* Filamin A is mutated in X-linked chronic idiopathic intestinal pseudo-obstruction with central nervous system involvement. *Am. J. Hum. Genet.* **80**, 751–758 (2007).
- Deglicenti, A. *et al.* A novel locus for syndromic chronic idiopathic intestinal pseudo-obstruction maps to chromosome 8q23-q24. *Eur. J. Hum. Genet.* **15**, 889–897 (2007).
- Siva, N. 1000 Genomes project. *Nat. Biotechnol.* **26**, 256 (2008).
- Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
- Seelow, D., Schuelke, M., Hildebrandt, F. & Nürnberg, P. HomozygosityMapper—an interactive approach to homozygosity mapping. *Nucleic Acids Res.* **37**, W593–W599 (2009).
- Bouchard, G., Roy, R., Casgrain, B. & Hubert, M. Population files and database management: the BALSAC database and the INGRES/INGRID system. *Hist. Mes.* **4**, 39–57 (1989).
- Thompson, E.A. Identity by descent: variation in meiosis, across genomes, and in populations. *Genetics* **194**, 301–326 (2013).
- Gramley, F. *et al.* Atrial fibrosis and atrial fibrillation: the role of the TGF- β 1 signaling pathway. *Int. J. Cardiol.* **143**, 405–413 (2010).
- Fu, M., Lui, V.C.H., Sham, M.H., Pachnis, V. & Tam, P.K.H. Sonic hedgehog regulates the proliferation, differentiation, and migration of enteric neural crest cells in gut. *J. Cell Biol.* **166**, 673–684 (2004).
- Sukegawa, A. *et al.* The concentric structure of the developing gut is regulated by Sonic hedgehog derived from endodermal epithelium. *Development* **127**, 1971–1980 (2000).
- Tessadori, F. *et al.* Identification and functional characterization of cardiac pacemaker cells in zebrafish. *PLoS ONE* **7**, e47644 (2012).
- Pater, E. *et al.* Distinct phases of cardiomyocyte differentiation regulate growth of the zebrafish heart. *Development* **136**, 1633–1641 (2009).
- Kline, A.D. *et al.* Cornelia de Lange syndrome: clinical review, diagnostic and scoring systems, and anticipatory guidance. *Am. J. Med. Genet. A* **143A**, 1287–1296 (2007).
- Gillis, L.A. *et al.* *NIPBL* mutational analysis in 120 individuals with Cornelia de Lange syndrome and evaluation of genotype-phenotype correlations. *Am. J. Hum. Genet.* **75**, 610–623 (2004).
- Deardorff, M.A. *et al.* *RAD21* mutations cause a human cohesinopathy. *Am. J. Hum. Genet.* **90**, 1014–1027 (2012).
- Deardorff, M.A. *et al.* *HDAC8* mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle. *Nature* **489**, 313–317 (2012).
- Musio, A. *et al.* X-linked Cornelia de Lange syndrome owing to *SMC1L1* mutations. *Nat. Genet.* **38**, 528–530 (2006).
- van der Lelij, P. *et al.* Warsaw breakage syndrome, a cohesinopathy associated with mutations in the XPD helicase family member DDX11/ChIR1. *Am. J. Hum. Genet.* **86**, 262–266 (2010).
- Caburet, S. *et al.* Mutant cohesin in premature ovarian failure. *N. Engl. J. Med.* **370**, 943–949 (2014).
- Yamagishi, Y., Sakuno, T., Shimura, M. & Watanabe, Y. Heterochromatin links to centromeric protection by recruiting shugoshin. *Nature* **455**, 251–255 (2008).
- Heijman, J., Dewenter, M., El-Armouche, A. & Dobrev, D. Function and regulation of serine/threonine phosphatases in the healthy and diseased heart. *J. Mol. Cell. Cardiol.* **64**, 90–98 (2013).
- Schmidt, D. *et al.* A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Res.* **20**, 578–588 (2010).

ONLINE METHODS

Study subjects. The ethical committees at all participating institutions approved the study, and all participants gave informed consent. Cases were ascertained through participating centers and underwent repeated clinical, electrocardiographic, echocardiographic and gastrointestinal assessments as part of their routine medical management. The diagnostic criterion for SSS was chronic, inappropriate sinus bradycardia associated with inappropriate junctional rhythm, sinus pauses and atrial dysrhythmias (frequent atrial premature, atrial tachycardia, atrial flutter and atrial fibrillation). A diagnosis of CIPO was made on the basis of mechanical obstruction of the intestine in the absence of an anatomical cause and evidence of impaired motility. Full-thickness biopsy material for two CAID cases and an unaffected control was available after surgery. Distribution of phenotypes in the CIPO validation cohort (total $n = 43$) was as follows: neurogenic, $n = 10$; myogenic, $n = 4$; mixed origin, $n = 2$; uncertain origin, $n = 27$.

Genetic analyses. Whole-exome sequencing. DNA was extracted from peripheral blood using the Qiagen Gentra PureGene Blood kit. We performed whole-exome sequencing on three cases (subjects 2, 8 and 12; **Fig. 1**) from apparently unrelated families to identify the causative mutation, under the hypothesis that CAID was an autosomal recessive trait caused by a rare allele with a heterozygous frequency of $<1\%$. The FORGE Canada team at the Michael Smith Genome Sciences Centre (Vancouver, British Columbia, Canada) used standard protocols for the Agilent SureSelect 50Mb exome enrichment kit with subsequent 100-bp paired-end sequencing on the Illumina HiSeq 2000 platform. Reads remaining after quality control were aligned to the reference human genome (hg18) using the Burrows-Wheeler Aligner, and variants were called using SAMtools 0.1.13 utilities^{31,32}. At least 96% of the targeted exomes were covered by 20-fold for each sample. Only variants (single-nucleotide variants and small insertion-deletions (indels)) inside protein-coding regions defined by Ensembl release 54 gene models were retained for downstream analysis³³. To obtain a tractable set of potentially causative candidates, we correlated variants with the dbSNP 129 and 130 databases, the 1000 Genomes Project database¹¹, an in-house database of normal germline variants developed at the Michael Smith Genome Sciences Centre and a similar database at the CHU Sainte-Justine Research Center that contains data from 360 French-Canadian subjects without a history of SSS or CIPO. Variant positions were annotated with reference to Build 36 of the NCBI genome assembly and sequences [NM_001199252](#) and [NP_001186184](#).

SGOL1 mutation genotyping. We performed PCR-based bidirectional Sanger sequencing to confirm the *SGOL1* mutations detected by whole-exome sequencing as well as to screen additional probands with either CAID or SSS alone. For dideoxy sequencing of *SGOL1*, PCR amplification products were bidirectionally sequenced using BigDye 3.1 chemistry on an ABI 3130xl Genetic Analyzer (Applied Biosystems). Primer sequences are provided in **Supplementary Table 4a**. Sequencing traces were aligned to the genomic reference sequence using Lasergene SeqMan Pro software (DNASTAR).

High-density genotyping. SNP data were generated at the McGill Genome Innovation Centre. Hybridization and laser scanning of Illumina HumanOmni5-Quad arrays with a mean call rate of 99.8% was performed according to the specifications of the manufacturer, with subsequent analysis on the GenomeStudio platform (Illumina). To perform homozygosity mapping, we analyzed SNP haplotype data for homozygous regions with HomozygosityMapper and PLINK^{12,13}. Genetic quality control included identity-by-descent testing and principal-components analyses. Relatedness testing was performed with PLINK, reporting pairs of individuals with genome identity (PLINK π -hat) > 0.2 as being closely related or cryptic duplicates. Haplotypes and their frequencies were estimated using PLINK and the Hapview algorithm for populations from the HapMap phase 3 and 1000 Genomes Project data sets^{11,34,35}. The populations in these data sets included ones of African, admixed American, East Asian, South Asian and European ancestry. In addition, allele frequencies from Alfred were used to obtain accurate and well-defined estimates of the frequency distributions of variants across multiple populations³⁶. The haplotype age for the French-Canadian founder population was estimated on the basis of haplotype alignments and a demographic model that was described previously^{37,38}, which takes population growth into account. CNV calls were generated with the PennCNV and CNVtools software package, using the log

R ratio (LRR) and B allele frequency (BAF) measures automatically computed by the BeadStudio software (Illumina)^{39,40}.

Genealogy reconstruction. Genealogies were reconstructed through the BALSAC Project at the Université du Québec in Chicoutimi, which holds an electronic database of the vital records for the French-Canadian population since 1608 (ref. 14). The gene dropping method⁴¹ implemented in the R GenLib software package (see URLs) was used to estimate the genetic contribution of each common ancestor to *SGOL1* mutation carriers and 228 unrelated controls. For each common ancestor, the probability of allele transmission was determined on the basis of known genealogies. Permutations were repeated 100 million times for each candidate ancestor, and the allele transmission probability was assessed on the basis of the assumption that each affected individual inherits two copies of the allele and the allele is absent in controls. The final probability for each common ancestor was obtained by multiplying the probabilities across all affected individuals and controls. The probability of ancestor i being the most likely common ancestor was given on the basis of the observed genetic information G that was estimated as the denominator sums of all the probabilities over all candidate ancestors.

Cell culture assays. All cell cultures were routinely tested for the absence of mycoplasma infection. All assays were carried out between passages 3 and 12, with a matched number of passages in each assay.

Proliferation assays. Primary dermal fibroblasts from normal controls and homozygous carriers of the *SGOL1* mutation encoding p.Lys23Glu were cultured in DMEM culture medium in the presence of 10% FBS and seeded at 4,000 cells/well in a 96-well plate in triplicate in three sets of plates. Cell proliferation was evaluated using WST-1 reagent (Roche) according to the manufacturer's instructions. The day after seeding, WST-1 was added to one set of plates to determine the background control. The other plates were maintained in culture for 24 and 48 h, and WST-1 was added at the end of each culture period. Proliferation was assessed by measuring the absorbance of the samples relative to the background controls at 440 nm.

Cell cycling and mitosis. To analyze cell cycling and mitosis, fibroblasts were cultured as described above and then washed with PBS and fixed with cold 70% ethanol for 1 h on ice. After fixation, cells were resuspended in PBS with 0.5% Triton X-100 and incubated at room temperature for 30 min. Cells were washed in PBS and incubated with Alexa Fluor 488-conjugated antibody to phosphorylated histone H3 (Cell Signaling Technology; 1:20 dilution in PBS with 1% BSA; 16-h incubation). Cells were washed and resuspended in PBS containing 150 μ g of RNase A (Fermentas) and 7.5 μ g of propidium iodide and incubated for 30 min at room temperature. Fluorescence was measured using a flow cytometer (BD Biosciences FACS Fortessa) and FlowJo software (Tree Star). Three cell lines were studied for each genotype in triplicate, with a minimum of 100,000 total events.

Immunofluorescence microscopy. Fibroblasts were cultured on poly-L-lysine-coated coverslips in 12-well plates and fixed for 5 min in 4% paraformaldehyde in PBS, permeabilized in PBS containing 0.25% Triton X-100 (PBST buffer), blocked for 1 h at room temperature in PBST buffer containing 5% goat serum and incubated overnight at 4 °C with primary antibody (1:90 dilution; mouse antibody to Sgol1, Abcam, ab58023). Cells were washed three times for 5 min per wash with PBS and incubated for 1 h at room temperature with Alexa Fluor 488-conjugated goat anti-rabbit secondary antibody (1:1,000 dilution; Molecular Probes/Invitrogen). Hoechst 33258 (1:500 dilution) was added for the final 5 min before cells were washed with PBS and mounted in ProLong antifade reagent (Molecular Probes/Invitrogen). Images were acquired with Velocity version 6.0 software (Improvision/PerkinElmer) controlling an Ultraview Vox spinning-disc confocal system (PerkinElmer), employing a CSU-X1 scanning unit (Yokogawa) and an Orca-R2 charge-coupled device camera (Hamamatsu) fitted to a Leica DMI6000B inverted microscope equipped with a 1.4-NA Plan Apochromat 100 \times oil immersion objective.

Image analysis. Images were exported from Velocity as "item as tiff" files and imported into ImageJ software (version 1.47). Using single confocal sections, regions of interest were manually drawn around each nucleus (using the Hoechst channel), and *SGOL1* pixel intensity values were extracted using the Sixteenbit Histogram extension plug-in. Pixel intensity values were plotted as a function of their frequencies (normalized by pixel numbers) using Prism software (GraphPad). *SGOL1* intensity distributions are

shown as the mean frequencies (\pm s.e.m.) for a total of 12 cells from 3 different homozygous cases or wild-type controls. The corresponding background distribution represents the data from similarly sized regions of interest selected from cell-free areas of the images.

Senescence assays. Senescence was measured with SA- β -gal staining at pH 5.5 using previously published protocols⁴². Briefly, 40,000 fibroblasts/well were seeded and grown in 6-well plates as described above and were analyzed at the indicated passages. All assays were performed in quadruplicate in three different cell lines for each wild-type control and homozygous carrier of the *SGOL1* mutation encoding p.Lys23Glu. After the aspiration of medium, cells were washed twice with PBS, fixed with 4% formaldehyde (Sigma) for 5 min at room temperature, washed twice with 2 ml of PBS and incubated for 16 h in a 37 °C chamber with staining solution at pH 5.5 (final concentrations: citric acid, 20 mM; phosphate, 40 mM; potassium ferricyanide, 5 mM; potassium ferrocyanide, 5 mM; NaCl, 150 mM; MgCl₂, 2 mM; X-gal in dimethylformamide, 1 mg/ml). Stained cells were counted manually for 200 cells, and numbers were documented digitally.

Flow cytometric TGF- β assays. Fibroblasts were grown to approximately 80% confluence in DMEM supplemented with 10% FBS in a 6-well plate. Cultures were then starved in 2% serum for 24 h before stimulation with vehicle or 10 ng/ml TGF- β 1 (R&D) for the indicated times. At each time point, cells were quickly lifted from the plates by adding trypsin and pipetting forcefully with a P1000 pipette; signaling was stopped by quickly mixing one volume of cells plus trypsin with one volume of 4% paraformaldehyde (EMS) in PBS. Cells were fixed for 10 min at 37 °C, washed once in PBS containing 0.5% BSA (flow cytometry staining buffer), and resuspended and incubated for 30 min in ice-cold 90% methanol. Cells were then washed once with flow cytometry staining buffer and stained at room temperature for 1 h with phycoerythrin-conjugated antibody to phosphorylated Smad2 and Smad3 (clone O72-670, BD Biosciences) by adding 5 μ l of antibody to 1×10^6 cells in a 100- μ l volume. Unbound antibody was removed by washing the cells twice in flow cytometry staining buffer. Stained cell suspensions were analyzed on a BD FACSVersa flow cytometer (BD Biosciences). Results were evaluated, including median fluorescent intensity calculation, using FlowJo software (Tree Star).

C banding. Metaphase chromosome spreads were prepared from peripheral blood samples, and Giemsa C banding was performed using standard cytogenetic protocols.

Immunohistochemistry on intestinal sections. Paraffin-embedded sections (8- μ m thick) were immunostained as previously described^{43,44}. Antigen retrieval was achieved by heating at 98 °C for 20 min in 1 mM citrate (pH 6.0). For *SGOL1* immunostaining, a mouse monoclonal antibody (ab58023, Abcam) was used at a 1:200 dilution as the primary antibody. In addition, antibodies to the neuronal marker TUJ1 (MMS-435P, mouse monoclonal antibody, Covance; 1:400 dilution), the TMEM16A marker for interstitial cells of Cajal (ab53212, rabbit polyclonal antibody, Abcam; 1:200 dilution), α SMA (IR611, Dako; prediluted), CD117 (A4502, Dako; 1:500 dilution) and CD3 (IR503, Dako; prediluted) were used. Trichrome-Masson staining was performed according to standard protocols (AR173, Dako). Immunohistochemistry control experiments were performed by excluding the primary antibody (data not shown).

Zebrafish studies. *Zebrafish husbandry and in situ hybridization.* The fish used in this study were kept under standard conditions as previously described⁴⁵. All zebrafish work conformed to ethical guidelines and was approved by the

relevant local ethics committees at the University of Utrecht. The *isl1*-mutant line *isl1K88X SA0029* has been described previously²⁰. *In situ* hybridization on zebrafish embryos was carried out as previously described⁴⁶. *In situ* hybridization on adult heart tissue was carried out as previously described with minor modifications⁴⁷. The sample size was not predetermined.

Heart rate measurements. Embryos at 3 d.p.f. were embedded in 0.25% agarose prepared in E3 embryonic medium with 16 mg/ml 3-aminobenzoic acid ethyl ester. Heartbeat was imaged using high-speed video recording at 28 °C and analyzed as previously described¹⁹. Statistical analysis and drawing of the box-whisker plot were carried out in Excel 2007 (Microsoft).

Antisense morpholino-mediated knockdown. Antisense morpholinos targeting *sgol1* were used at a final concentration of 500 μ M in a volume of 1 nl per embryo and injected into embryos at the one-cell stage (**Supplementary Table 4b**). Antisense morpholinos targeting the zebrafish *SCN5A* orthologs *scn12aa* and *scn12ab* and *tp53* have been described previously and were used at final concentrations of 200 μ M for *scn12aa* and 100 μ M for *scn12ab* and *tp53* (ref. 48). Uninjected embryos from the same egg lay were used as controls for all experiments.

31. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
32. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
33. Flicek, P. *et al.* Ensembl 2012. *Nucleic Acids Res.* **40**, D84–D90 (2012).
34. International HapMap Consortium. The International HapMap Project. *Nature* **426**, 789–796 (2003).
35. Gabriel, S.B. *et al.* The structure of haplotype blocks in the human genome. *Science* **296**, 2225–2229 (2002).
36. Rajeevan, H. *et al.* ALFRED: an allele frequency database for microevolutionary studies. *Evol. Bioinform. Online* **1**, 1–10 (2005).
37. Labuda, M. *et al.* Linkage disequilibrium analysis in young populations: pseudo-vitamin D-deficiency rickets and the founder effect in French Canadians. *Am. J. Hum. Genet.* **59**, 633–643 (1996).
38. Labuda, D., Zietkiewicz, E. & Labuda, M. The genetic clock and the age of the founder effect in growing populations: a lesson from French Canadians and Ashkenazim. *Am. J. Hum. Genet.* **61**, 768–771 (1997).
39. Glessner, J.T., Li, J. & Hakonarson, H. ParseCNV integrative copy number variation association software with quality tracking. *Nucleic Acids Res.* **41**, e64 (2013).
40. Subirana, I., Diaz-Uriarte, R., Lucas, G. & Gonzalez, J.R. CNVassoc: association analysis of CNV data using R. *BMC Med. Genomics* **4**, 47 (2011).
41. MacCluer, J.W., VandeBerg, J.L., Read, B. & Ryder, O.A. Pedigree analysis by computer simulation. *Zoo Biol.* **5**, 147–160 (1986).
42. Itahana, K., Campisi, J. & Dimri, G.P. Methods to detect biomarkers of cellular senescence: the senescence-associated β -galactosidase assay. *Methods Mol. Biol.* **371**, 21–31 (2007).
43. Notarnicola, C. *et al.* The RNA-binding protein RBPM2 regulates development of gastrointestinal smooth muscle. *Gastroenterology* **143**, 687–697 (2012).
44. Rouleau, C., Matécki, S., Kalfa, N., Costes, V. & de Santa Barbara, P. Activation of MAP kinase (ERK1/2) in human neonatal colonic enteric nervous system. *Neurogastroenterol. Motil.* **21**, 207–214 (2009).
45. Westerfield, M. *The Zebrafish Book: A Guide for the Laboratory Use of Zebrafish Danio (Brachydanio rerio)* (University of Oregon Press, Eugene, Oregon, USA, 2000).
46. Thisse, C. & Thisse, B. High-resolution *in situ* hybridization to whole-mount zebrafish embryos. *Nat. Protoc.* **3**, 59–69 (2008).
47. Moorman, A.F., Houweling, A.C., de Boer, P.A. & Christoffels, V.M. Sensitive nonradioactive detection of mRNA in tissue sections: novel application of the whole-mount *in situ* hybridization protocol. *J. Histochem. Cytochem.* **49**, 1–8 (2001).
48. Chopra, S.S. *et al.* Voltage-gated sodium channels are required for heart development in zebrafish. *Circ. Res.* **106**, 1342–1350 (2010).