Supplemental Appendix

This appendix has been provided by the authors to give readers additional information about their work.

CONTENTS

Case Histories

Patient 1 (P1) was born at term to non related, healthy Welsh parents with a weight of 1.26kg (-5.4SD), head circumference 27.6 cm (-5.6SD) and length 39cm (-6.0SD). She has four unaffected siblings. She has shown lifelong severe linear growth retardation (adult height 1.07m (-9.4SD); weight 21.6 kg (-9.7 SD) but no evidence of osteodysplasia. She has facial dysmorphism including a small lower jaw and a prominent midface, however her secondary teeth are small but normally formed. Endocrine evaluation at 5 years old showed normal generation of both IGF1 and growth hormone. Mild global learning difficulty was formally diagnosed at 3.5 years old (at which time developmental age by Griffiths Assessment was 2.75 years). Primary gonadal failure was diagnosed based on failure to exhibit signs of puberty by 14 years, with elevated gonadotrophins (LH 41.7 mU/l and FSH 71.7 mU/l), undetectable oestrogen, and prepubertal appearance of ovaries and uterus, leading to treatment with oestrogen and later progesterone.

The karyotype was 46, XX with no evidence of increased sister chromatid exchange on diagnostic testing for Bloom's syndrome at 1 year old. At 5 years old, analysis of 200 unsynchronised lymphocytes in culture revealed 0.36 cytogenetic aberrations per cell (0.08 in 200 control cells), comprising 16% chromosome gaps, breaks or acentric fragments (2.5% in controls), of which 7.5% were acentric fragments (1% in controls). 0.09 chromatid aberrations per cell were seen (0.07 in controls). At 11.8 years old, unsynchronized lymphocyte cultures revealed a persistent but lesser degree of chromosome damage, with 5% chromosome gaps and 4% chromatid gaps seen in 100 cells compared to 0.66% and 2% respectively in 150 control cells. No acentric fragments were seen at this age. At 16.9 years old, no evidence of increased basal or mitomycin C (MMC)-provoked chromosome breakage was seen using up to 0.5 μg/ml of MMC. All cytogenetic evaluations were undertaken in the clinically accredited Cytogenetic Laboratory for Wales, University Hospital of Wales, Cardiff, UK.

At 4 years old, a mildly distended abdomen with elevated fasting serum triglyceride concentration (3.8 mmol/l (336 mg/dl)(normal < 1.8 mmol/l) after 12 h of fasting), and elevation of serum alanine and aspartate transaminase were recorded, consistent with development of fatty liver. At 11 years old acanthosis nigricans, a cutaneous marker of insulin resistance, was reported, and impaired glucose tolerance with extreme elevation of fasting plasma insulin to 4,396 pmol/l (normal <60 pmol/l) despite a normal blood glucose of 4.4 mmol/l (79 mg/dl) was recorded. At 12 years old, serum was noted incidentally to be lipemic, suggesting excess chylomicrons, and fasting serum triglyceride concentrations were elevated at 4.8 mmol/l (424 mg/dl) with a concomitantly low HDL cholesterol of 0.47 mmol/l (18 mg/dl)(normal 0.7-1.8 mmol/l). At 13 years old, symptomatic diabetes developed and metformin was commenced, initially with good effect. Pioglitazone and then insulin were started at 16 years old, with the dose of insulin rising to 12 units/kg/day by 17 years old. At that stage, prior to a previously reported 3 month open label uncontrolled trial of a subcutaneous IGF1/IGFBP3 composite (Regan FM *et al*, J Clin Endocrinol Metab. 95(5):2113-22, 2010), detailed metabolic evaluation revealed poor metabolic control (HbA1c 10.4%; serum triglyceride 25.5 mmol/l (2,257 mg/dl), and fatty liver (liver triglyceride content

22.6% by magnetic resonance spectroscopy; serum AST and ALT raised). Pioglitazone was increased to 45mg once daily, however poor glycemic control persisted despite insulin doses up to 24 units/kg/day. Glycemic control improved significantly from 22 to 24 years old, when HbA1c fell to around 6.0%. At 23 years old, pioglitazone was stopped because of accumulation of subcutaneous adipose tissue on the limbs and abdomen and an increase in weight to 23.5 kg.

On evaluation at 24 years old, weight was 21.6 kg (-9.0SD), and height 1.07m (-9.4SD). There was no clinical evidence of premature ageing nor malignancy, however severe acanthosis nigricans persisted. Therapy consisted of sitaglitpin 50 mg once daily, simvastatin 10 mg once daily, hormone replacement therapy, and 3.7 units/kg/day of insulin. Non fasting plasma insulin was 3,030 pmol/l with concomitant C peptide 357 pmol/l (normal 170-830 pmol/l), leptin 15.1 μg/l (normal 2.4-24.4 μg/l), and adiponectin 1.4 mg/l (normal 4.4-17.7 mg/l).

Patient 2 (P2) was born at term to non-related French parents, with birthweight 1.2 kg (-5.6SD), and length 39 cm (-6.0SD). She was congenitally blind, reportedly due to bilateral retinal detachment, although records of the primary clinical evaluation were unavailable. Early development was normal and mild global learning difficulties were observed although the patient developed well given her level of physical handicap. No further clinical record was available until 11.5 years old when, on assessment for a febrile illness, glycosuria was noted. Fasting blood glucose was 6.1 mmol/l (110 mg/dl), and normal glucose excursion was noted on oral glucose tolerance testing, although plasma insulin was reported to be extremely high. Fasting glucose levels were subsequently found to lie persistently between 5.5 and 11.1 mmol/l (100-200 mg/dl), with associated hypertriglyceridemia up to 11.9 mmol/l (1,053 mg/dl)).

Episodes of acute pancreatitis at 12.5 and 13.8 years old were attributed to severe hypertriglyceridemia and produced pancreatic exocrine insufficiency that was treated with oral enzyme replacement. Despite a low fat diet severe hypertriglyceridemia persisted, with occasional appearance of eruptive xanthomata. Peak recorded weight of 10.1kg (-10.7 SD) was achieved at 13 years old. No evidence of puberty was apparent by 17 years old. MRI revealed normal brain and pituitary gland structure, but a pre-pubertal uterus and undetectable ovaries. By 20 years old, height was 0.97 m (-11.0 SD), weight 8.17 kg (-26.5 SD), and head circumference 41 cm (-10.4 SD). A skeletal survey revealed gracile long bones, widened metaphyses, and some cone-shaped epiphyses.

By 22 years old, there was progressive, symptomatic hyperglycemia with concomitant weight loss, with weight of 7.5 kg. By 25 years old, weight was 7.3 kg. At that stage the HbA1c 11.1% with a fasting plasma insulin of 70 pmol/l, C-peptide 650 pmol/l, triglyceride 50 mmol/l (4,425 mg/dl), and elevated serum AST and ALT. Under hyperinsulinemic euglycemic clamp conditions only 4mg glucose/kg/minute were required intravenously to maintain euglycemia in the face of intravenous infusion of 10 mU/kg/minute of insulin (corresponding to 2000-3000 pmol/l plasma insulin). Collectively these findings demonstrated 1. Severe insulin resistance 2. relative insulin deficiency and 3. very poorly controlled

diabetes. Subcutaneous insulin was commenced and titrated to high doses, peaking at 67 units/kg/day without hypoglycemia.

At 26 years old, clinical features of heart failure developed and echocardiography demonstrated left ventricular hypertrophy, attributed to hypertension, for which nifedipine was commenced. At 32 years old weight was 11.8 kg (-19.2SD), HbA1c was 8.0%, and triglyceride, AST and ALT were persistently raised. P2 died suddenly at 33 years old presumably due to the rupture of a previously identified abdominal aortic aneurysm. No autopsy was undertaken.

Evidence Against Pathogenicity of WDR36 and RUSC1 variants

WDR36 has previously been studied in the context of human disease and linked with human primary open-angle glaucoma, with some studies, but not all, supporting this link (Monemi *et al*, Hum. Molec. Genet. 14: 725-733, 2005; Fingert *et al,* Arch. Ophthal. 125: 434-436, 2007; Pasutto *et al,* Invest. Ophthal. Vis. Sci. 49: 270-274, 2008; Frezzotti *et al,* Brit. J. Ophthal. 95: 624-626, 2011). Irrespective of whether or not this gene has a role in glaucoma, these studies show that the D658G variant identified in Patient 1 has been observed in both glaucoma cases and unaffected controls across a range of studies (Monemi *et al* 2005; Fingert *et al* 2007, Pasutto *et al* 2008) with no other associated adverse phenotypic presentation associated with this variant. Furthermore, a published zebrafish knockdown of *Wdr36* demonstrates that it functions in nucleolar processing of 18S rRNA, which is required for ribosome biogenesis (Skarie *et al* Hum. Molec. Genet. 17: 2474-2485, 2008). The reported mutant is from a zebrafish line with a viral insert in the wdr36 gene, and the authors show that heterozygous mutants do not show any phenotypic differences from wild-type, and that homozygous mutants have small heads and small eyes but appear normal in length, a finding corroborated by morpholino knockdown in zebrafish. Detailed histological examination showed several ocular abnormalities not observed in our patient, as well as other morphological abnormalities including liver necrosis and absence of swim bladder inflation. Histology of gut structures showed severe differentiation defects in the liver, pancreas and intestine. None of these features overlap the phenotype of our patient, suggesting that loss of function of *WDR36* leads to a different phenotype from that which we report here.

RUSC1 mutations can be excluded as the cause of the observed phenotype with reasonable confidence based on human genetic data. Analysis of sequence data from 4,300 European American and 2,203 African American NHLBI Exomes identified 51 additional *RUSC1* variants that are predicted to alter protein sequence with consequences which are equally or more deleterious than those of the variants reported in Patient 1. Simulations were undertaken in which 100,000 pairs of parental haplotypes were sampled, with the probability of having any of the 53 mutations being based on their population frequency in either the European American or African American samples. It was thereby estimated that 1.13% of the European American and 4.88% African American offspring would be compound heterozygous for likely loss-of-function mutations. Given this much higher predicted prevalence of this disease than occurs in reality, and in particular given the much higher expected predicted prevalence in African Americans compared to samples of European ancestry, the *RUSC1* variants identified in Patient 1 are very unlikely to be causal.

Supplemental Materials and Methods

SMC5-6 expression analysis.
Urea-based whole cell extracts were immunoblotted for SMC5 and SMC6 using anti-SMC5 (epitope; AA181-305) and anti-SMC6 (epitope; AA103-636) rabbit polyclonal antibodies generated by Elaine Taylor and Alan Lehmann (Sussex, UK). Blots were reprobed with α-tubulin to confirm loading. For chromatin fractionation, 10^7 cells were re-suspended in 100 ml of *hypotonic* buffer (10 mM HEPES pH 7.5, 1.5 mM $MgCl_2$, 1 mM DTT, 10 mM NaF, 1 mM $NaVO_3$, 10 mM β-glycerolphosphate, 0.5% IGEPAL (NP-40) and protease inhibitor cocktail (Promega)), and incubated on ice for 15 min, pelleted, washed in hypotonic buffer before being re-suspended in *hypertonic* extraction buffer (hypotonic buffer supplemented with 0.5 M NaCl). Samples were incubated for 15 min before being extracted with urea-based extraction buffer and sonicated. Blots were re-probed for the heterochromatin marker KAP1 (Bethyl) to confirm loading.

Assessment of Stability of Mutant NSMCE2/MMS21 proteins
Total RNA extracted from fibroblasts and LCLs according to the standard protocol for the RNEasy kit (Qiagen) was used for MMLV (Promega) cDNA synthesis. Quantitative real-time PCR analysis of NSMCE2/MMS21 mRNA expression was undertaken using a gene-specific ABI TaqMan Gene Expression Assay (Hs00329126_m1) in ABI TaqMan PCR Mastermix and analyzed on a ABI 7900 HT detection system. Results were normalized to levels of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) mRNA. Sequencing of NSMCE2/MMS21 in fibroblast-derived cDNA was done using the BigDye Terminator v3.1 Cycle Sequencing Kit (ABI) and primer 5'-TGCTACATTGGATCGGCAAC-3' or 5'-CAGCTTGAAGACCACATTGCT-3' or 5'-GAAGAGGACGCCATTGTTCG-3'.

For protein stability experiments, HeLa cells were transfected with Myc-tagged human NSMCE2/MMS21 expression vectors using Lipofectamine 2000 before treatment at 20 h post-transfection with protein synthesis inhibitor cycloheximide (CHX) at 0.1 mg/ml. Western blotting of cleared whole cell extracts was undertaken using anti-Myc and actin antibodies (both from Millipore) and the relative levels of Myc-tagged protein and actin were then determined by densitometry. The fraction of normalized Myc-tagged protein remaining after 26 hours of CHX treatment was calculated.

Viability and apoptosis analysis of LCLs.
Viability analysis was conducted using the CellTitre-Blue assay (Promega) according to the manufacturer's instructions and analysed using a GloMax microplate reader (Promega). For apoptosis analysis, urea-based whole cell extracts were blotted using an antibody specific to the p85 form of PARP (Promega G734A).

Evaluation of mRNA and Protein Levels of NSMCE2/MMS21 Species Following Microinjection of mRNA into Zebrafish Embryos.
For each datapoint a pool of 10 embryos was used to generate cDNA for the qRT-PCR reaction, each run in duplicate. RNA for GFP (50 pg) was co-injected with NSMCE2 RNA (100 pg) and quantification of GFP transcript levels were used to normalize NSMCE2/MMS21 transcript abundance. Fluorescent microscopy was used to determine the level of protein expression of GFP-tagged WT and variant NSMCE2/MMS21 in RNA injected embryos compared to non-injected embryos. Fluorescent images were captured using a fluorescent stereomicroscope (MZ FLIII, Leica) coupled with Nikon D3 digital camera. For each time point and image the exposure time and the gain were the same.

Supplemental Tables

	Primer Sequence
NSMCE2-ex1	(F) GTCCTCTTCTATCCGCAGTCC (R) ATCGCTAATTCTCACTGCAATCTAC
NSMCE2-ex2/3	(F) TTTCTAACCCATTTATCTTGTAGGC (R) CCACACAGTGTTACCTCAAAGTAGA
NSMCE2-ex4	(F) TTGACTGATGTAGATTGTATTGATGC (R) TTTGGCAATAGATGCATATTAAAGTT
NSMCE2-ex5	(F) GAATAACCACTTGTGAGTGTGAGAG (R) CTGTTCATCAAGTCAGTTCCCTAAT
NSMCE2-ex6	(F) GTTTCTATGCATCCAGTCCTTACAT (R) ATGATCAACTTCCAGTAGCTGTTTC
NSMCE2-ex7	(F) TGGTTTGATTTCACTTAGGGAGTTA (R) GAACGGAATTTAGAAGAGACATTGA
NSMCE2-ex8	(F) ATAAGGCAGGCACTACTCACATATC (R) TGATTGCATTTCTGTTGTGTAACTT

Supplemental Table 1: Primer sequences for Sanger sequencing of *NSMCE2* exons and intron/exon boundaries.

Gene	Mutation	Chr	Position (GRCh37)	ref	alt	1000 Genomes				NHLBI Exomes		UK10K	ODEX	dbSNP
						EUR	AMR	AFR	ASN	EA	AA	Cohorts	Exomes	v137
NSMCE2	**S116Lfs*18**	8	126194424	T	-	a	a	a	a	0.0002	0.0000	a	a	a
NSMCE2	**A234Efs*4**	8	126379079	-	AGGG	a	a	a	a	a	a	a	a	a
RUSC1	**S171A**	1	155292075	T	G	0.000	0.063	0.003	0.000	0.0014	0.0578	0.0019	a	rs74118445
RUSC1	**S614F**	1	155295681	C	T	0.004	0.000	0.003	0.000	0.0017	0.0002	0.0033	0.0024	rs148123009
WDR36	**A163V**	5	110434448	C	T	0.003	0.000	0.000	0.000	0.0057	0.0009	0.0093	0.0060	rs62376783
WDR36	**D658G**	5	110454719	A	G	0.004	0.000	0.003	0.000	0.0072	0.0020	0.0073	0.0072	rs34595252

Supplemental Table 2: Compound heterozygous mutations found in Patient P1. Chromosome positions are relative to NCBI build 37 and dbSNP identifiers to version 137. Ref refers to the NCBI reference sequence allele, alt refers to the alternate allele. Non-reference allele frequencies are given for a) 1000 Genomes (release 2012-07-19) – EUR (European ancestry), AMR (from the Americas), AFR (West African ancestry), ASI (East Asian ancestry); b) NHLBI Exome Variant Server (release ESP6500SI-V2, Apr-2013) – EA (European American ancestry), AA (African American ancestry) and c) 3,781 Genomes from the UK10K Cohorts group (UK10K Cohorts, release 2012-06-02) - all of European ancestry and d) 409 CoLaus Exomes (ODEX, release Dec-2010) – all of European ancestry. Variants absent from the control populations are labelled “a”.

	Alleles	rs12542833	rs3860845	rs12682006	rs7836506	rs7831515	rs6470338	rs7813225	p.S116Lfs*11	rs10091027	rs7008482	rs6992463	rs4242367	rs10755926	rs7004739	rs10100121	rs2293890	p.A234Efs*4	rs10094181	rs10956242	rs4871590	rs2954015	rs2954006	Haplotype frequency
	0	C	G	C	C	C	T	A	del	A	T	A	G	G	G	A	G	AGGG	T	C	G	G	C	
	1	A	A	A	T	T	A	G	T	G	G	G	A	A	A	G	C	del	G	T	A	T	G	
P1 & P2 trio haplotypes	P1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1.79%
		0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.79%
	Father	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	Mother	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	
	P2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1.79%
		0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.79%
	Father	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	0	1	1	1	1	0	
	Mother	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	
		1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Individual haplotypes for 54 ethnically matched controls		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	13.39%
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11.61%
		0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	1	1	1	0	7.14%
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	5.36%
		0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	1	1	0	2.68%
		0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	1	1	0	0	2.68%
		0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	1	1	0	1	2.68%
		0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1.79%
		0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	0	0	1	0	1.79%
		0	0	0	0	0	0	0	0	0	1	1	1	1	0	1	1	0	1	0	0	1	1	1.79%
		0	0	0	0	0	0	0	0	0	1	1	0	0	1	1	1	0	1	1	1	1	0	1.79%
		0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	1	0	0	0	1	1.79%
		1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	1	1	1	0	1.79%
		1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	0	1	0	0	0	0	1.79%
		1	1	1	0	1	1	1	0	0	0	0	1	1	1	1	1	0	1	1	1	0	1	1.79%
		1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	0	1	0	0	0	1	1.79%
		1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1.79%
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0.89%
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0.89%
		0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0.89%
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0.89%
		0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	1	0.89%
		1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.89%
		1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0.89%
		1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0.89%
		0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	1	0.89%
		1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0.89%
		1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0.89%
		1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0.89%
		1	1	0	0	1	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0.89%
		0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	1	0	0	0.89%
		0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	1	0	1	0.89%
		1	1	1	0	1	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0.89%
		1	1	0	0	1	0	0	0	1	1	1	0	0	0	0	0	0	0	1	0	0	1	0.89%
		0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	1	1	1	1	1	0.89%
		1	1	1	0	1	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0.89%
		0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	1	1	1	1	0.89%
		0	0	0	0	1	0	0	0	1	1	1	0	0	1	1	1	0	1	1	1	1	0	0.89%
		1	1	0	0	1	1	1	0	1	0	0	0	0	1	1	1	0	1	1	1	0	0	0.89%
		1	1	0	0	1	0	0	0	1	1	1	1	1	1	1	1	0	1	0	0	0	0	0.89%
		1	1	1	0	1	1	1	0	0	0	0	1	1	1	1	1	0	1	0	0	0	0	0.89%
		1	1	1	0	1	1	1	0	0	0	0	1	1	1	1	1	0	1	0	0	1	0	0.89%
		0	0	0	0	1	1	1	0	1	1	1	0	0	1	1	1	0	1	1	1	1	0	0.89%
		1	1	0	0	1	1	1	0	1	0	0	0	0	1	1	1	0	1	1	1	0	1	0.89%
		1	1	0	0	1	0	0	0	1	1	1	1	1	1	1	1	0	1	0	0	0	1	0.89%
		0	0	0	0	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1	1	0	0.89%
		1	1	1	0	1	1	1	0	0	0	0	1	1	1	1	1	0	1	1	1	0	0	0.89%
		1	1	0	0	1	0	0	0	1	1	1	1	1	1	1	1	0	1	1	1	0	1	0.89%
		1	1	0	0	1	1	1	0	1	1	1	1	1	1	1	1	0	1	0	0	0	1	0.89%
		1	1	0	0	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1	1	0	0.89%
		1	1	0	0	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0.89%
		1	1	0	0	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0.89%

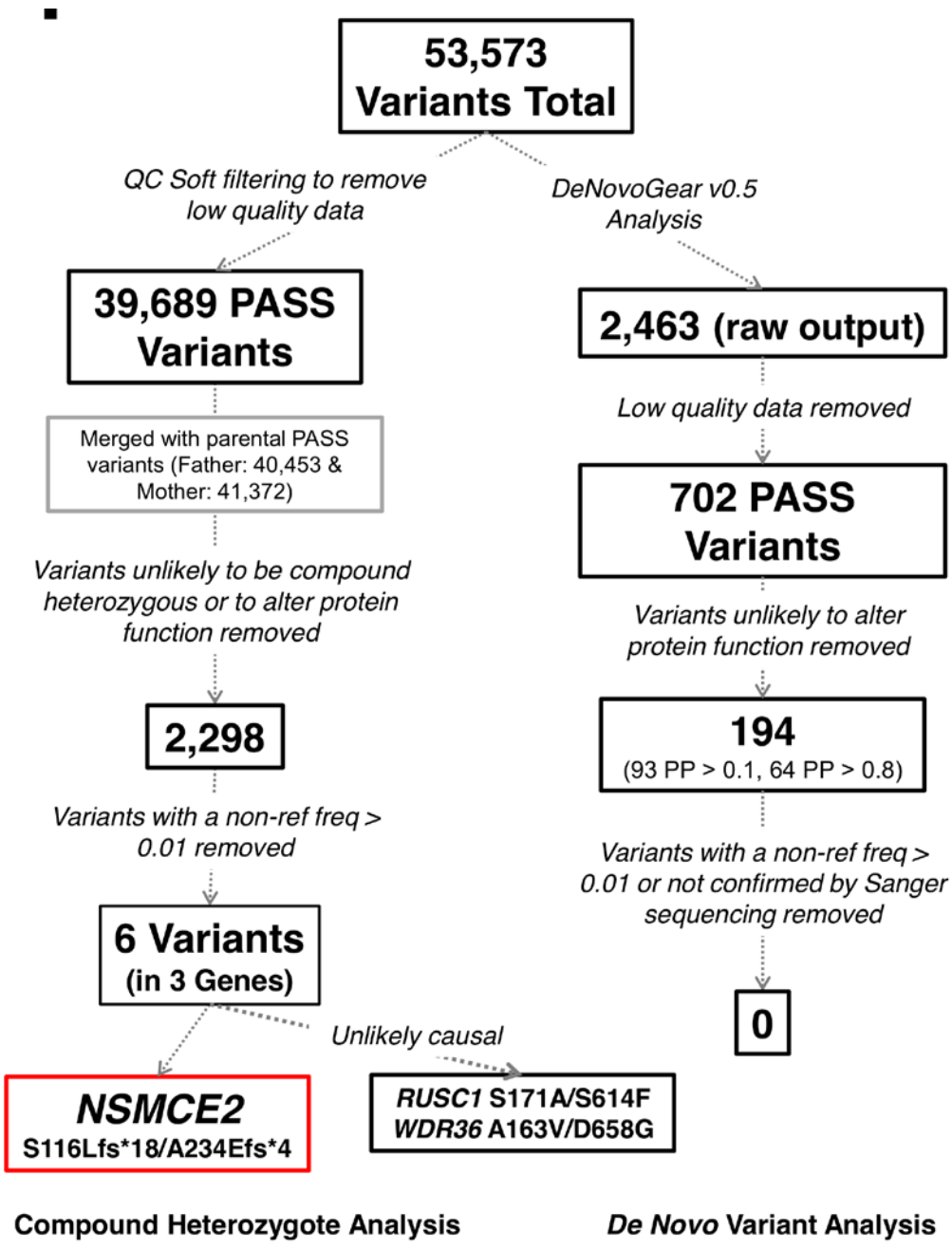
Supplemental Table 3: Haplotype Analysis of Affected Families and Controls. (Legend overleaf)

Supplemental Table 3: Haplotype Analysis of Affected Families and Controls. Haplotypes around the p.Ser116Leufs*18 and p.Ala234Glufs*4 mutations in the P1 and P2 trios and 54 ethnically matched controls. Haplotype frequencies were calculated in patients and controls, ignoring additional family members. The two haplotypes shared by P1 and P2 are indicated in blue and yellow in parents and offspring. The p.S116Lfs*18 and p.A234Efs*4 mutations are highlighted in boldface. P1: Patient 1, P2: Patient 2

	Morpholino Sequence	**Target**
Nsmce2 sp1	GCTAAACGCATATTTCTCACCTTGT	Exon1-intron1
Nsmce2 sp2	TGGTCTTTTGGGTTTCCAACCTTGC	Exon3-intron3
Std Control Mo	CTCTTACCTCAGTTACAATTTATA	N/A

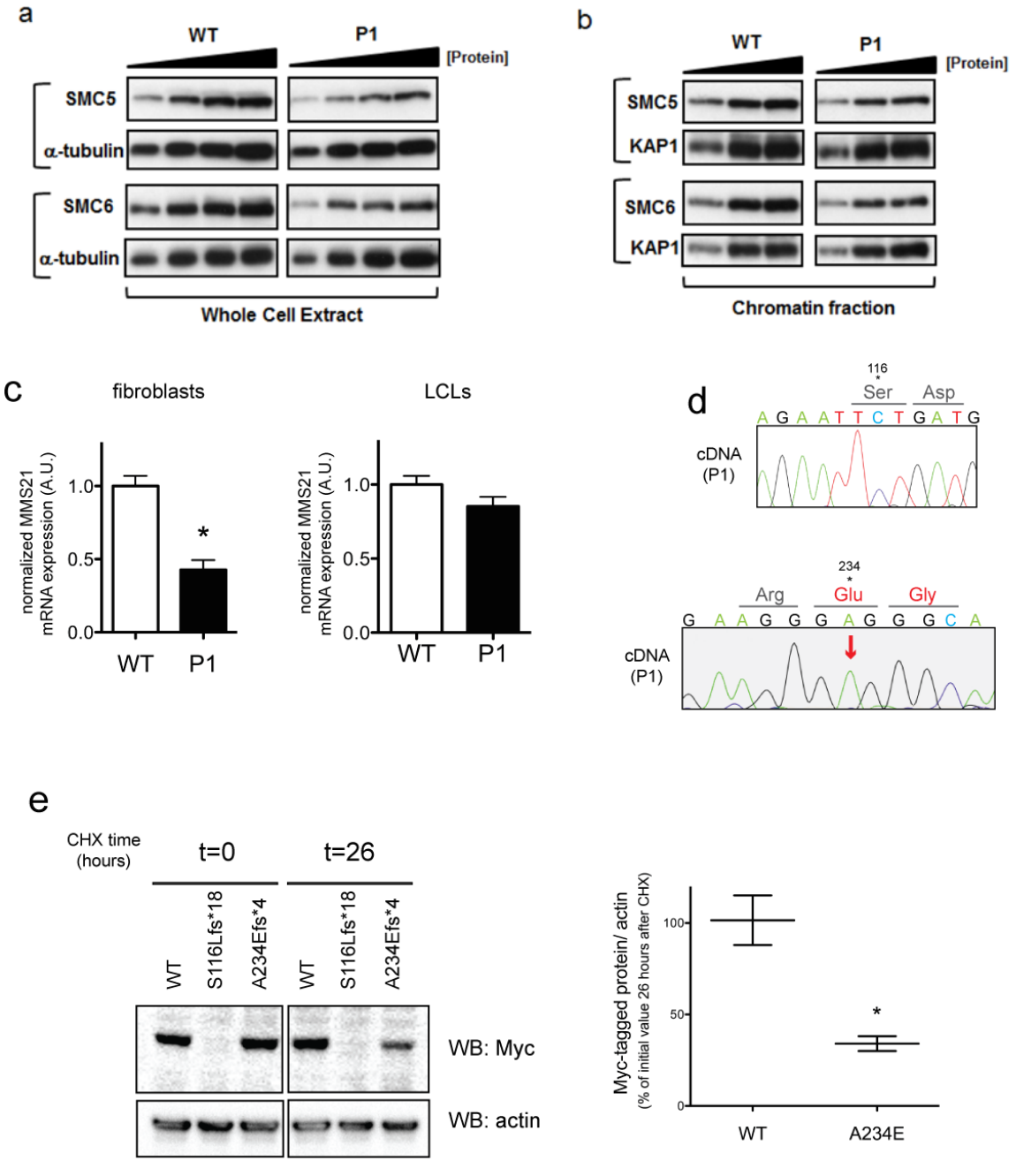
	Primer Sequence
Sp1	(F) GCTGGGAATAGTCGAAGCAC (R) CTGTGATTTCGCAGGTCTGA
Sp2	(F) TGAAGAAACTGGGGGAGATG (R) GAGAGGGCAAATGAAGTTCG
Control (Agpat2)	(F) GACTCGGTTCAGAGCCTACG (R)TTGTGCAGCGATCTGGTAAG

Supplemental Table 4: Primer and Morpholino Sequences used in Zebrafish Studies



Supplemental Figure 1: Schematic of exome sequence data analysis. (Legend overleaf)

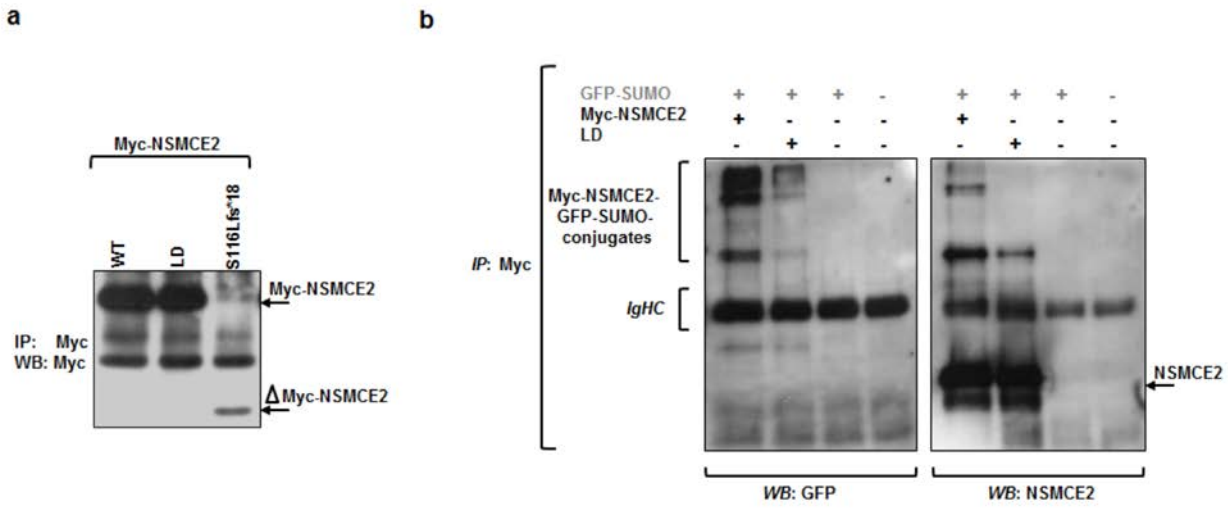
Supplemental Figure 1: Schematic of the analysis of the exome-wide sequence data. Overview of the filtering procedure for compound heterozygous mutations on the left and for potential *de novo* variants on the right (see Methods). Analysis of compound heterozygous mutations by filtering for genes with at least two rare, heterozygous mutations predicted to change protein structure identified 6 variants in 3 genes, four of which were present in control populations (albeit with a non-reference allele frequency < 0.01) and in genes unlikely to contribute to the phenotype of the patient. All variants detected by *de novo* analysis were either also found in control populations (with a non-reference ("non ref") allele frequency > 0.01), or proved not to be true mutations when re-sequenced using Sanger sequencing. PP represents Posterior probability – a measure of the likelihood that the mutation is *de novo.* DeNovoGear recommend using a cutoff PP of 0.8.



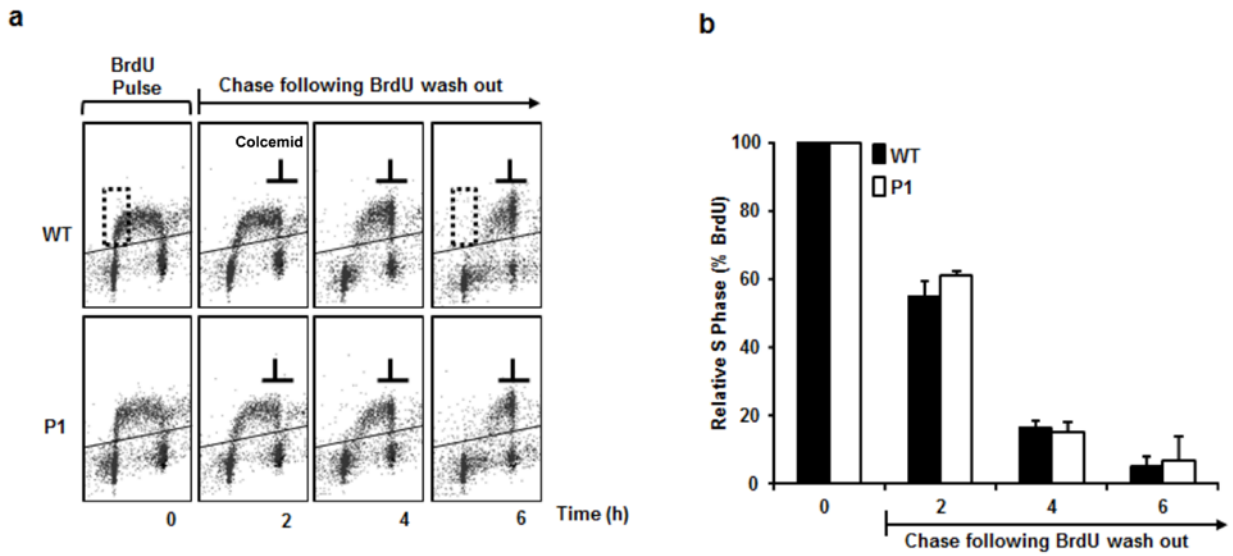
Supplemental Figure 2: Instability of Mutant NSMCE2/MMS21 (Legend overleaf)

Supplemental Figure 2: Instability of Mutant NSMCE2/MMS21

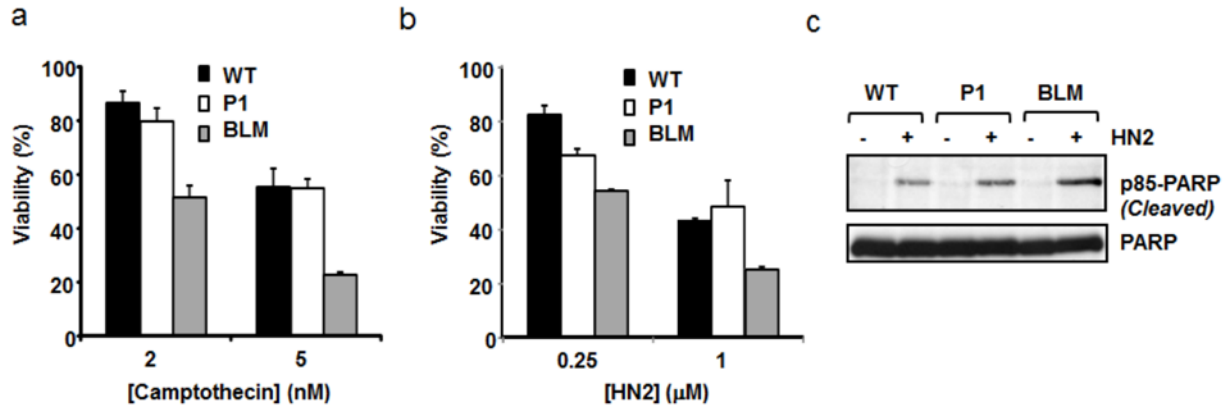
a) Western blot analysis of SMC5 and SMC6 levels in wild-type (WT) and P1-derived LCLs using urea-based whole cell extraction. Protein levels were titrated as indicated (5-50 μg). b) SMC5 and SMC6 levels were also investigated in WT and P1-derived LCLs following chromatin fractionation. Again protein extract were titrated as indicated. c) Quantitative real-time PCR analysis of NSMCE2/MMS21 mRNA expression in P1 fibroblasts or LCLs (black bars) and cells from healthy control subjects (WT, white bars) normalized to mRNA expression of glyceraldehyde-3-phosphate dehydrogenase (GAPDH). d) Sequencing of NSMCE2/MMS21 in P1 fibroblast-derived cDNA (S116Lfs region chromatogram on white background; A234Efs region chromatogram on gray background). e) Stability of myc-tagged NSMCE2/MM21 species in HeLa cells after inhibition of protein synthesis using cycloheximide. Cells were transfected with Myc-tagged human wild-type MMS21 (WT), p.Ser116Leufs*18 MMS21 (S116Lfs*18), or p.Ala234Glufs*4 MMS21 (A234Efs*4), and treated 20 h post-transfection with protein synthesis inhibitor cycloheximide (CHX) (0.1 mg/ml) before immunoblotting of cleared whole cell extracts. Relative levels of Myc-tagged protein and actin were then determined by densitometry. Plotted data represent the fraction of normalized Myc-tagged protein remaining after 26 h of CHX treatment (t_0 - t_{26}) (right panel). Numerical data are represented as mean ± S.E.M. (n=3). * = $p<0.05$ (unpaired, 2 tailed *t*-test)



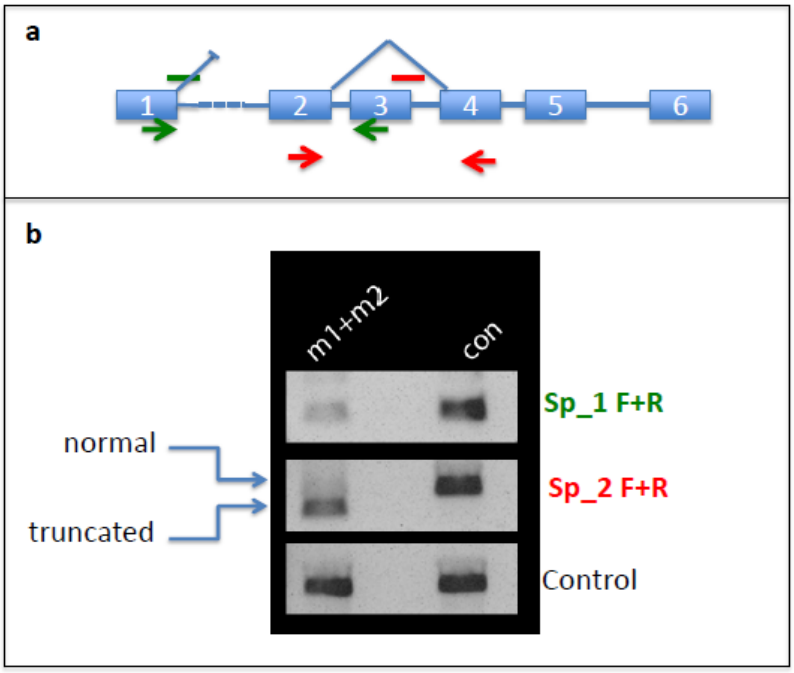
Supplemental Figure 3: Overexpression of myc-tagged NSMCE2 and Validation of Auto-SUMOylation Assay a) Relative size and amount of wild-type (WT), SUMO-ligase deficient (LD) and p.S116Lfs*18 Myc-NSMCE2 obtained following transient expression in A549 cells and immunoprecipitation (IP) using anti-Myc. **b)** Auto-SUMOylation assays of wild-type (WT) and SUMO-ligase deficient (LD) showing the SUMO-conjugated forms of NSMCE2 detected using either anti-GFP (left hand panel) or using anti-NSMCE2 (right hand panel). The slower migrating higher molecular weight SUMO conjugates are markedly reduced in LD compared to WT, as expected. IgHC; immunoglobin heavy chain.



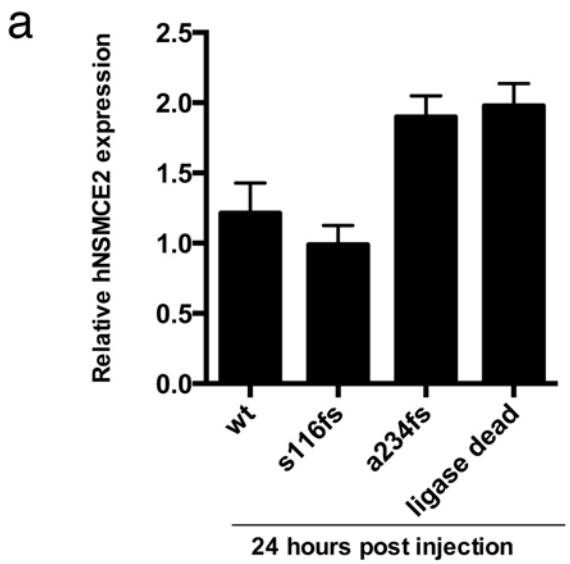
Supplemental Figure 4: S phase progression kinetics in P1-derived LCLs Unperturbed P1 LCLs show similar kinetics in S phase progression compared to those of wild-type (WT). Cells were pulsed with BrdU to label S phase, BrdU was removed and cells were chased in a Colcemid block for up to 6 h. The Colcemid treatment prevents transition out of mitosis. **a)** representative FACS profiles at each time point. **b)** The kinetics of S phase progression from early S phase (boxed area) was quantified over time. Numerical data are represented as mean $\pm$ S.D. (n=3).



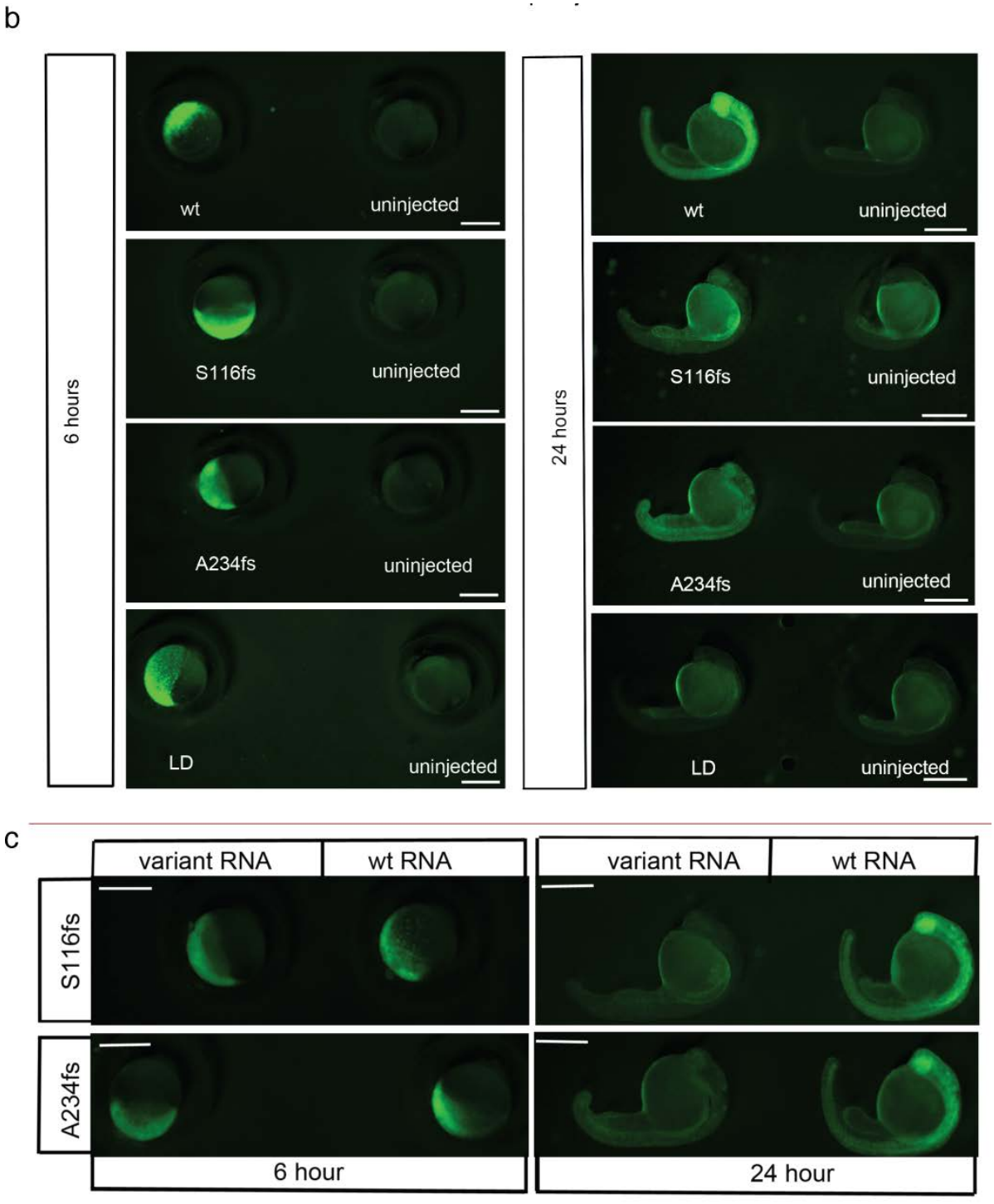
Supplemental Figure 5: Response of P1-derived LCLs to treatment with DNA Damaging agents. a) Camptothecin, via its role in Topoisomerase I (Topo I) inhibition, is a potent inducer of DNA breaks, particularly DNA double strand breaks when the inhibited and stabilized Topo I-DNA cleavable complex collides with an active replication and/or transcription fork. LCLs were treated with Camptothecin as indicated, and incubated for 3 days before being assessed for viability using the CellTitre Blue system. LCLs from a Bloom syndrome individual (GM09960) were used as a comparator. b) LCLs were treated as indicated with Mechlorethamine (HN2), a bi-functional nitrogen mustard which cross-links DNA via N7 of guanine. Cells were cultured for 48 h before assaying for viability. c) LCLs were treated with 2.5 μM HN2 and apoptosis levels investigated using PARP cleavage 24 h following treatment. Numerical data are represented as mean ±S.D. (n=3-5)



Supplemental Figure 6: Schematic of Nsmce2 zebrafish gene structure and knockdown strategy. a) Target sites of splice-blocking morpholinos. Red bar and arrows indicate the position of MO 1 and corresponding primers (sp 1F+R), green bar and arrows indicate the position of MO 2, target site and corresponding primers (sp2 F+R). **b)** Reverse transcriptase polymerase chain reaction (RT-PCR) analysis of nsmce2 mRNA transcripts in control and *nsmce2* MO-injected embryos.



Supplemental Figure 7: Evaluation of RNA and protein levels of wt and NSMCE2/MMS21 variants following microinjection of RNA into zebrafish embryos (legend overleaf)



Supplemental Figure 7: Evaluation of RNA and protein levels of wt and NSMCE2/MMS21 variants following microinjection of RNA into zebrafish embryos (legend overleaf)

Supplemental Figure 7: Evaluation of RNA and protein levels of WT and NSMCE2/MMS21 variants following microinjection of RNA into zebrafish embryos. a) Quantification of WT and variants of NSMCE2/MMS21 transcript levels in RNA-injected embryos by quantitative real time–PCR. RNA for GFP (50 pg) was co-injected with NSMCE2 RNA (100 pg) and quantifications of GFP transcript levels were used to normalize the abundance of NSMCE2/MMS21. b) Representative fluorescent micrographs showing protein expression of GFP-tagged WT and variants of NSMCE2/MMS21 in RNA-injected embryos compared to non-injected embryos (uninjected). c) Comparison of protein expression between WT and NSMCE2/MMS21 variants following RNA microinjection into zebrafish embryos. Scale bars, 0.5 mm. For each time point the exposure time and gain microscopy settings were the same for every acquired image. Numerical data are shown as mean ± S.E.M. (n=3)

UK10K Researchers

Principal Applicant

- Richard Durbin, Wellcome Trust Sanger Institute

Co-applicants

- Jeffrey Barrett, Wellcome Trust Sanger Institute
- Ines Barroso, Wellcome Trust Sanger Institute
- George Davey-Smith, University of Bristol
- Ismaa Sadaf Farooqi, University of Cambridge
- Matthew Hurles, Wellcome Trust Sanger Institute
- Stephen O'Rahilly, University of Cambridge
- Aarno Palotie, Wellcome Trust Sanger Institute
- Nicole Soranzo, Wellcome Trust Sanger Institute
- Tim Spector, King's College London
- Eleftheria Zeggini, Wellcome Trust Sanger Institute

Named collaborators

- Phil Beales, University College London
- Jamie Bentham, University of Oxford
- Shoumo Bhattacharya, University of Oxford
- Douglas Blackwood, Edinburgh University
- Patrick Bolton, King's College London
- Gerome Breen, King's College London
- Krishnan Chatterjee, University of Cambridge
- David Collier, King's College London
- David Fitzpatrick, Edinburgh University
- Louise Gallagher, Trinity College Dublin
- Daniel Geschwind, UCLA, USA
- Hugh Gurling, University College London
- Steve Humphries, University College London
- Peter McGuffin, King's College London
- Antony Monaco, University of Oxford
- Francesco Muntoni, University College London
- Michael Owen, Cardiff University
- Lucy Raymond, University of Cambridge
- David Savage, University of Cambridge
- Peter Scambler, University College London
- Robert Semple, University of Cambridge
- David Skuse, University College London
- David St Clair, University of Aberdeen
- Nic Timpson, Bristol University