The American Journal of Human Genetics, Volume *106*

Supplemental Data

De Novo Frameshift Variants in the Neuronal Splicing Factor *NOVA2* Result in a Common C-Terminal Extension and Cause a Severe Form of Neurodevelopmental Disorder

Francesca Mattioli, Gaelle Hayot, Nathalie Drouot, Bertrand Isidor, Jérémie Courraud, Maria-Victoria Hinckelmann, Frederic Tran Mau-Them, Chantal Sellier, Alica Goldman, Aida Telegrafi, Alicia Boughton, Candace Gamble, Sebastien Moutton, Angélique Quartier, Nolwenn Jean, Paul Van Ness, Sarah Grotto, Sophie Nambot, Ganka Douglas, Yue Cindy Si, Jamel Chelly, Zohra Shad, Elisabeth Kaplan, Richard Dineen, Christelle Golzio, Nicolas Charlet-Berguerand, Jean-Louis Mandel, and Amélie Piton

SUPPLEMENTARIES

Supplemental Note: Case Reports

Individual 1 was the second child of non-consanguineous union. The family history was unremarkable. The pregnancy was complicated by hydramnios. He was delivered at term and at birth he had a normal mensuration. He showed neonatal feeding difficulties. He had no dysmorphic features. He was not able to sit independently until 1 year of age and showed truncal hands stereotypic movements. First episode of seizure was noted at the age of 2 years and 6 months. He was noted to smile a lot having frequent unprovoked laughter and being attracted to water. At last evaluation at the age of 9 years and 6 months, he had no speech and could not walk. Physical examination showed lower limbs hypertonia with brisk reflexes. His height was 122cm (-2SD), weight 21kg (-2,5SD) and head circumference 50 cm (-2,5SD). Brain MRI showed cortical atrophy. Angelman syndrome was suspected but 15q11-13 methylation was normal and *UBE3A* direct sequencing did not identify any mutation. Myoclonic epilepsy was treated by sodium valproate.

Individual 2 was the first child of non-consanguineous union. The family history was unremarkable. The pregnancy was uncomplicated, with normal screening ultrasounds. He was delivered via vaginal delivery and mensuration at birth were normal. He was hypotonic without facial dysmorphism. He was not able to sit independently until 1 year of age and walk until 3,5 years. He smiled a lot and was hyperactive. Physical examination showed brisk lower limbs reflexes and ataxic gait. His height was 105 cm (-2SD), weight 16kg (-2,5SD) and head circumference 49,3 cm (-2,2SD). He showed stereotypic hand movements. Brain MRI was

normal. His behavior and frequent outbursts of anger suggested autism spectrum disorder. Angelman syndrome was suspected but 15q11-13 methylation study was normal.

Individual 3 was the first the fourth child of non-consanguineous union. Family history was negative for neurodevelopmental phenotypes. The pregnancy was uncomplicated and patient was born via C-section at 37 weeks gestation. Neonatal course was benign, although sagittal craniosynostosis was not early in infancy. Developmental concerns became apparent to family around 12 months of age, due to speech and gross motor delay. Neurologic evaluation at 2 years of age was notable for hypotonia, Chiari malformation type I, and global developmental delay. Brain MRI at 2 years of age identified stable Chiari I malformation and congenital absence of the left internal carotid artery. The patient had normal SNP microarray. At her last clinical evaluation at 4 years and 8 months of age, she had been clinically diagnosed with autism spectrum disorder. Physical examination showed microcephaly, ears with pointed helices, ataxia, and hypotonia. She had impaired social interaction with delayed speech for age. Growth parameters were as follows height 103cm (29%), weight 17.7kg (57%). Patient was benefitting from speech, occupational, and music therapy.

Individual 5 is a 22-year-old male with intellectual disability, drug-resistant epilepsy, global developmental delay, and autism. According to his mother, he was born from a 27-year-old mother and 33-year-old father following an uneventful pregnancy and delivery. His immediate postnatal development seemed normal. However, at about 8 months of age, mother felt that "something was wrong" as he was developing slower as compared to his older brother and cousins. He started walking independently at two years of age, babbled some words but never developed language. At the age of three, he was formally diagnosed with a global

developmental delay and autism was diagnosed at nine years of age around the time of the onset of his myoclonic astatic epilepsy. His most disabling seizure types are drop attacks occurring about 15 times a month. His second most common seizure type starts with a fast movement followed by rocking back and forth while remaining semi-responsive. Parents also reported occasional starring spells occurring on average once a month. The patient had a corrective surgery for strabism at 12 months. He is ambulatory but nonverbal and requires total care and assistance with all activities of daily living, sometimes he would use a spoon for eating. He lives with his family and attends an adult learning center during the daytime. There is no history of ID in the family. A gadolinium enhanced brain MRI performed at 17 years of age showed an absence of neuronal migration defects. It was remarkable only for a slight cerebellar volume loss, thinning of the an otherwise normally configured corpus callosum in keeping with a global white matter volume loss, most pronounced in biparietal regions. Due to frequent nocturnal awakenings, he was referred for a sleep study at 16 years of age. During the study, interictal discharges in the midline central region were noted but no electrographic or electro-clinical seizures were recorded. There were no central or obstructive apneas. While oxygenation was maintained at or above 96%, significant non-obstructive hypercapnia with pCO2 over 52mmHg was recorded. At the time of the sleep study, his height was 181.3 cm, weight 48.7 kg and the body mass index was 14.82 kg/(m^2). EKG lead documented a normal sinus rhythm throughout the study. He had EEG studies done as a child but the results were not available for review. His general examination was unremarkable. There was no apparent craniofacial dysmorphology and no abnormal cutaneous lesions. His neurological evaluation was remarkable for a poor eye contact and lack of language. He was awake and oriented to his parents, exhibiting behavioral stereotypies manifested by body rocking movements. Cranial nerve and motor examination was unremarkable, deep tendon reflexes were present throughout and pathological reflexes were absent. Sensory evaluation was intact to light touch otherwise not testable. Coordination testing

was limited by patient's inability to cooperate but there was no gross ataxia or dysmetria or appendicular tremor, gait was intact. Romberg was negative. Genetic testing performed upon transition to the Adult Epilepsy Clinic included chromosomal microarray and whole exome sequencing including mitochondrial sequence analysis.

Individual 6 was the second child of non-consanguineous union. The family history was unremarkable. The pregnancy was uncomplicated. Birth measurements were normal. Developmental delay was reported from the age of 6 months, with hypotonia. She was not able to sit independently until 1 year and 3 months and to walk until 3 years and 9 moths. At last evaluation at the age of 5 years and 5 months, she had no speech but could understand simple sentences. Feeding difficulties were reported with hypersensitivity to some textures. She showed mild ataxic gait and flapping hands stereotypic movements. Physical examination showed mild dysmorphic features with deeply set eyes, anteverted nares, and deeply grooved philtrum. Her height was 108 cm (0 SD), weight 16kg (-1SD) and head circumference 48 cm (-2SD). Lower limbs reflexes were normal. Brain MRI was normal at the age of 9 months. Angelman syndrome was suspected but 15q11-13 methylation study was normal.

Supplementary Tables

Table S1. List of genes differentially expressed (DE) after *NOVA2* inactivation in hNSCs

Table S2. List of the different splicing events occurring after *NOVA2* inactivation in hNSCs identified by LeafCutter program.

Supplementary Figures

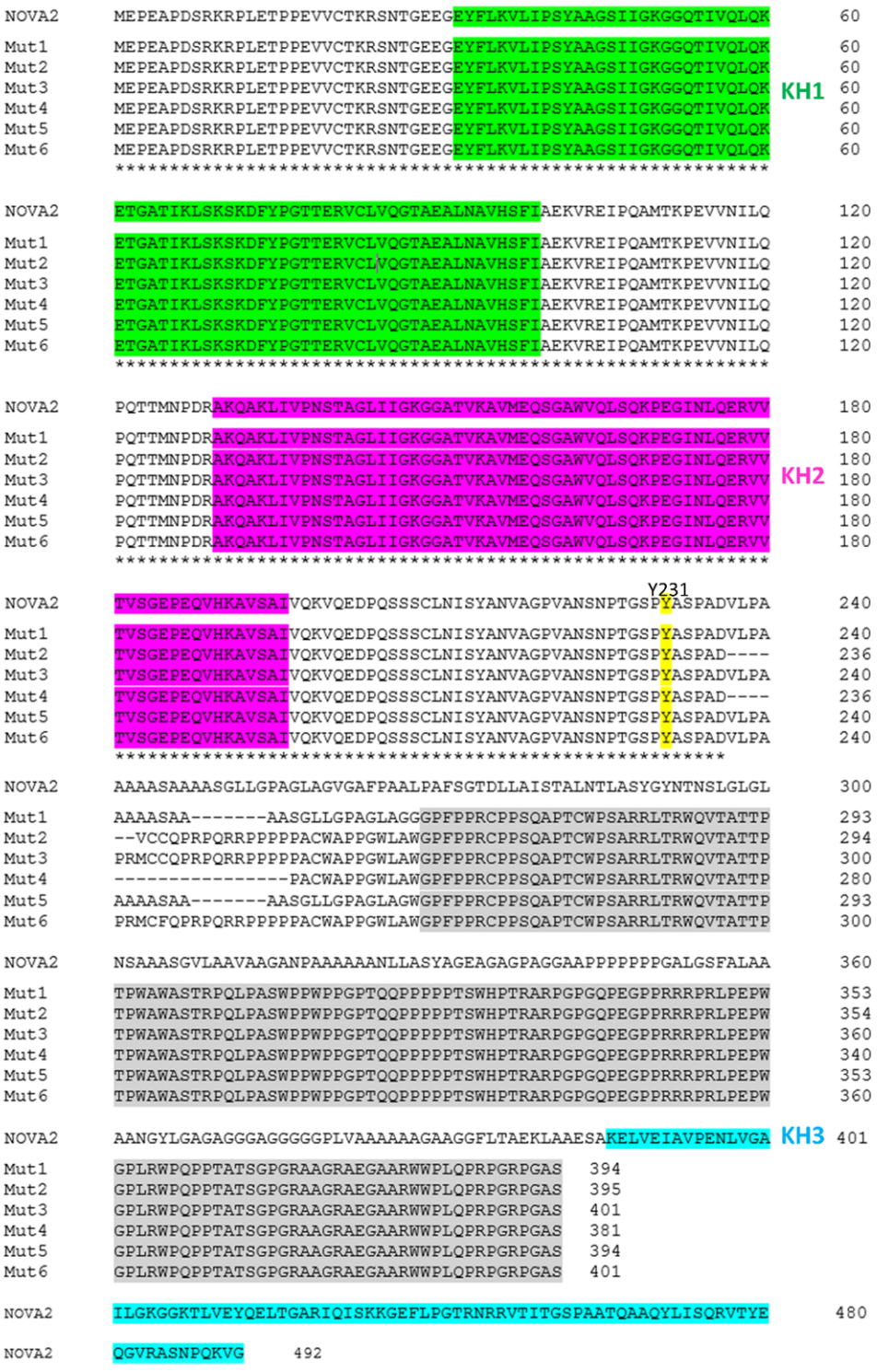

Figure S1. Alignement of wild type and variant NOVA2 proteins using ClustalW 2.1.

In grey, the common amino acid sequence introduced by all the frameshift variants. The KH domains (green for KH1, pink for KH2 and blue for KH3) and the position of the Tyr231 are represented. "*" indicates the amino acids common between the wild-type and variant proteins. Mut1 corresponds to c.782del, p.(Val261Glyfs*135) (identified in Individual 1), Mut2 to c.710_711dup, p.(Leu238Cysfs*159) (identified in Individual 2), Mut3 to c.701_720dup, p.(Ala241Profs*162) (identified in Individual 3), Mut4 to c.709_748del, p.(Val237Profs*146) (identified in Individual 4), Mut5 to c.781del, p.(Val261Trpfs*135) (identified in Individual 5), Mut6 to c.720_721insCCGCGGATGTGCTTCCAGCC, p.(Ala241Profs*162)(identified in Individual 6).

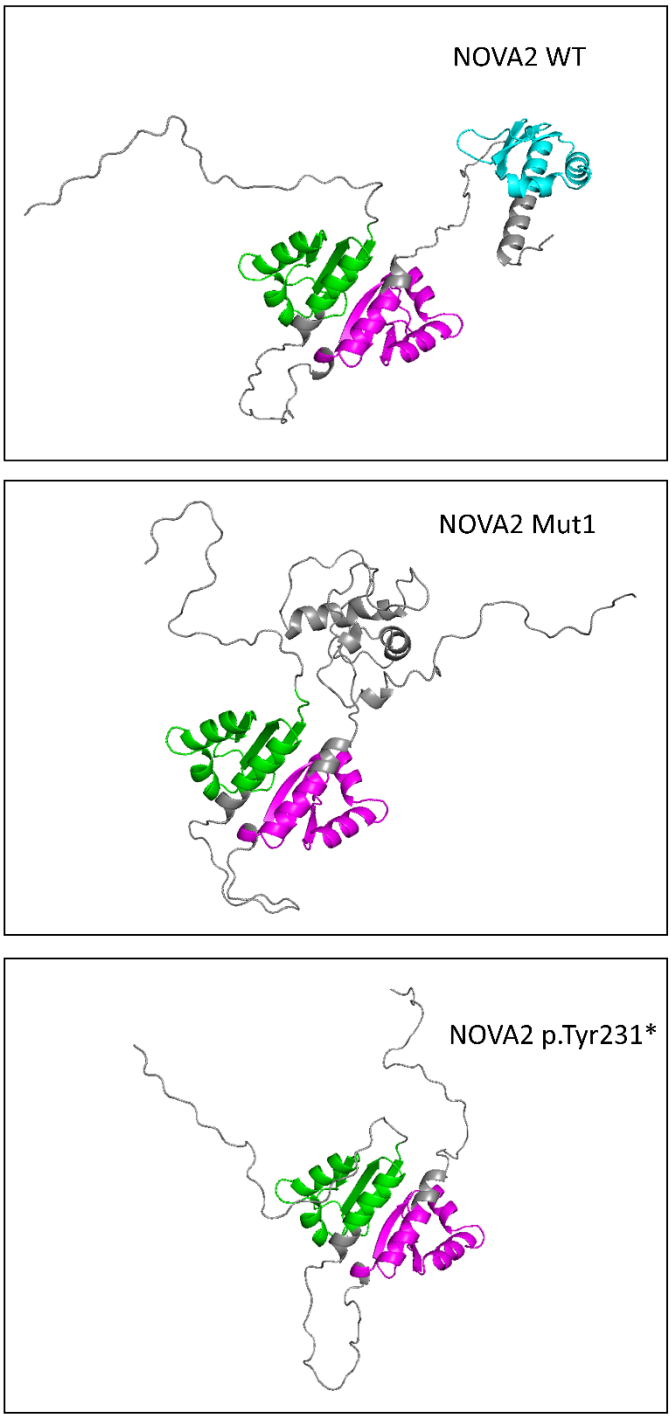

Figure S2. Predicted structures for wild-type and variant NOVA2 proteins

The 3D structures were generated from FASTA protein sequences using RaptorX for NOVA2 wild-type (WT), and mutant NOVA2 protein carrying patient's variant Mut1 or carrying a stop codon p.Tyr231*. Predicted alpha-helices and beta-sheets are represented. The three KH domains are indicated (green for KH1, pink for KH2 and blue for KH3).

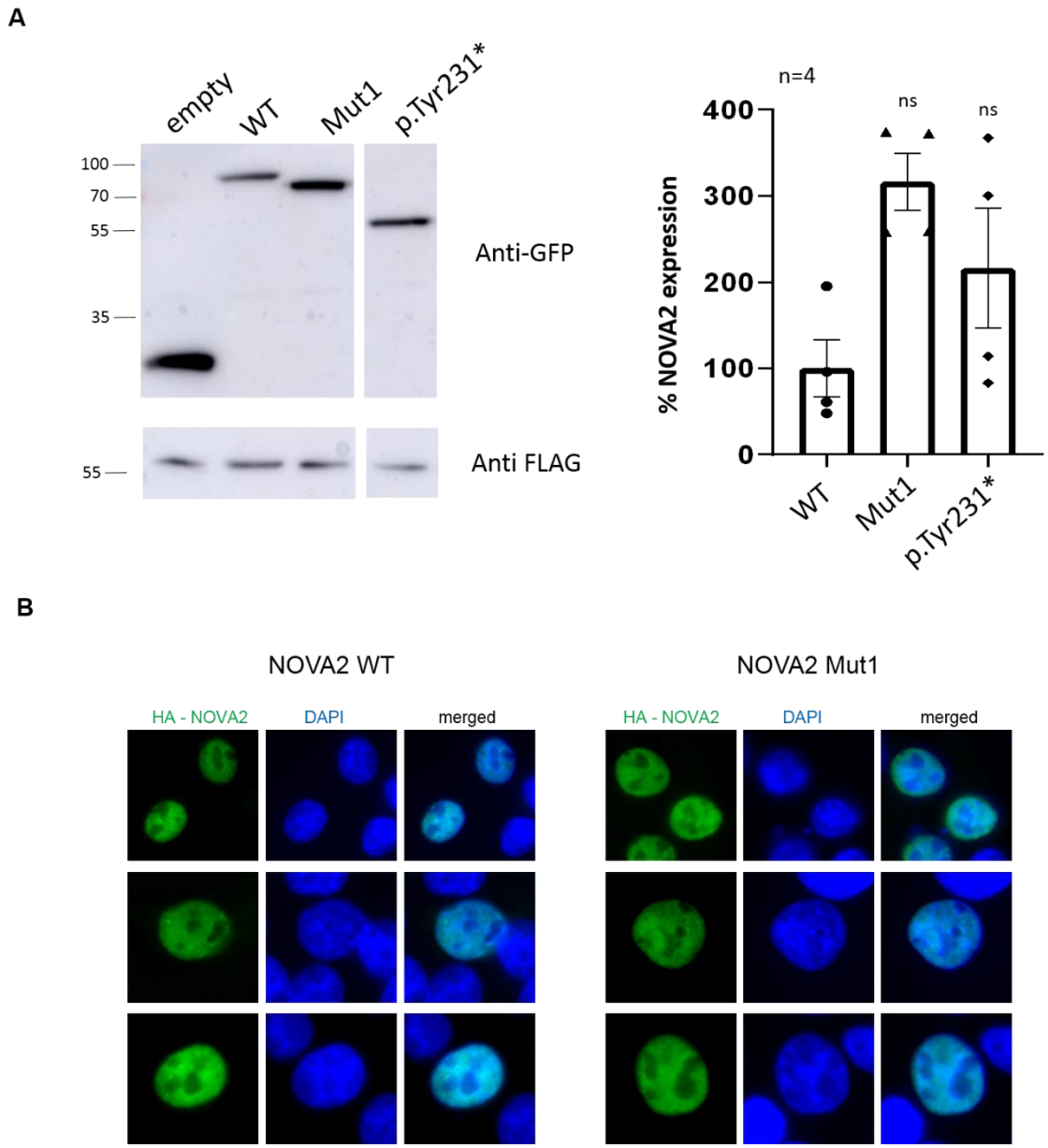

Figure S3. Expression and cellular localization of wild-type and Mut1 variant NOVA2 protein in HeLa cells

(A) Expression of NOVA2 in HeLa cells. HeLa cells were co-transfected with EGFP-tagged NOVA2 wild-type (WT), Mut1 or Tyr231* cDNA and a plasmid with a FLAG-tagged protein as a control for transfection. Cells were harvested 24 h after transfection, and expression of NOVA2 was analyzed by SDS–PAGE and immunoblotting with anti-GFP and anti-FLAG antibodies. Quantification of NOVA2 expression: the ratio between GFP and the transfection

control FLAG was calculated from four independent experiments and each condition was compared to the WT. The error bars indicate the standard error mean (SEM). Kruskal-Wallis ANOVA with Dunn's multiple comparisons, ns: non-significant (**B)** Cellular localization of NOVA2 proteins in Hela cells. Hela cell lines were transiently transfected with HA-NOVA2 WT or HA-NOVA2 variant Mut1 and immunofluorescence experiments using an anti-HA antibody revealed a nuclear localization of both WT and variant NOVA2 proteins. The DAPI (4′,6- diamidino- 2- phenylindole) staining indicates the position of the nuclei.

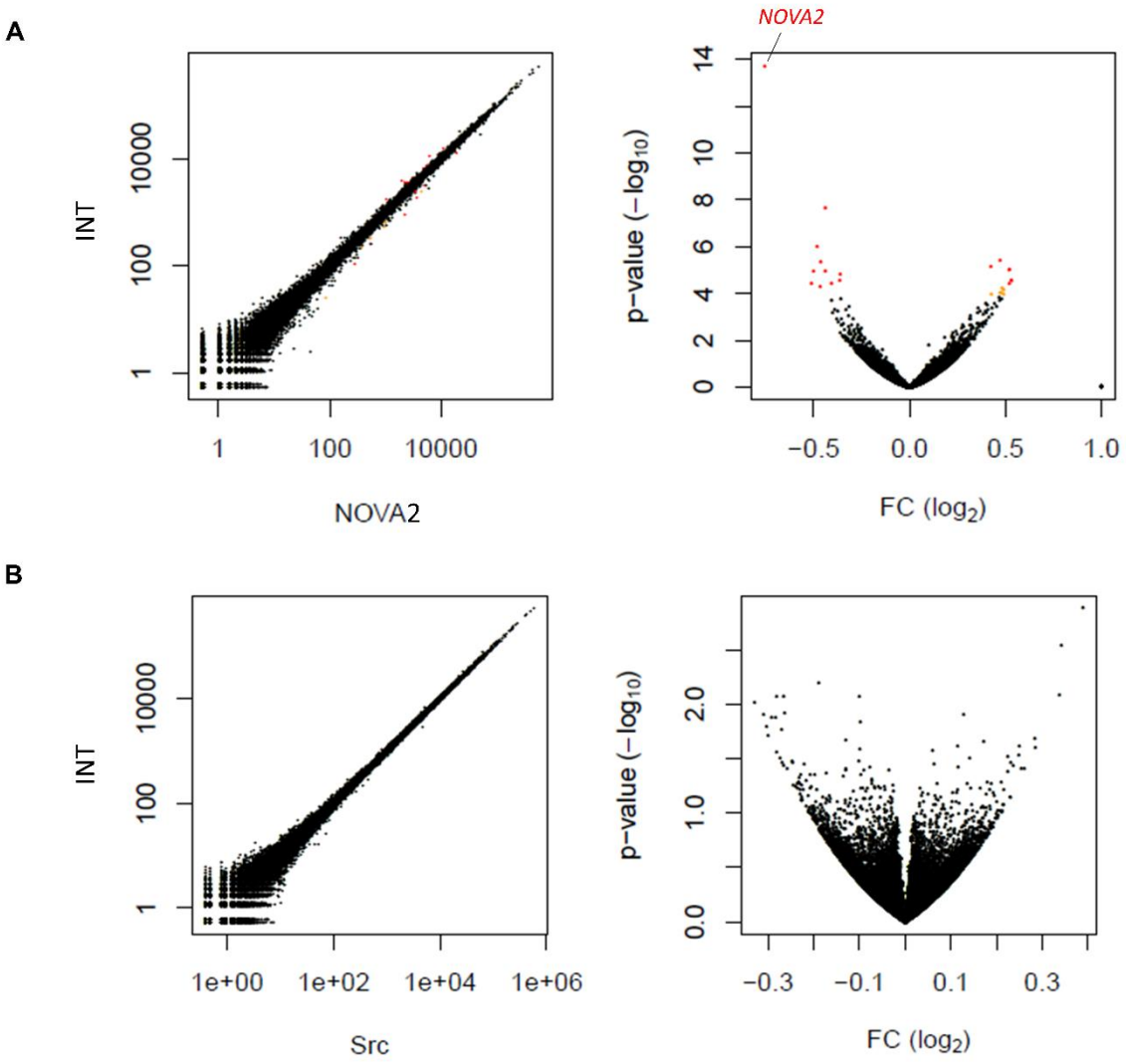

Figure S4. Volcano plot of differential expression after *NOVA2* inactivation in hNSCs

The graph on the left represents a scatterplot of the normalized gene expression. For a given gene, its value is obtained by mean on all samples belonging to the same condition. The graph on the right represents a volcano plot. The x-axis represents the log2 Fold Change values and the y-axis the p-values (in -log10). Red points show values with an adjusted p-value < 0.05 and orange points show values with an adjusted p-value < 0.1. Expression in (**A)** hNSCs treated with non-specific scramble siRNA compared to untreated cells (INT) and expression in (**B)** hNSCs treated with *NOVA2* scramble siRNA compared to untreated cells.

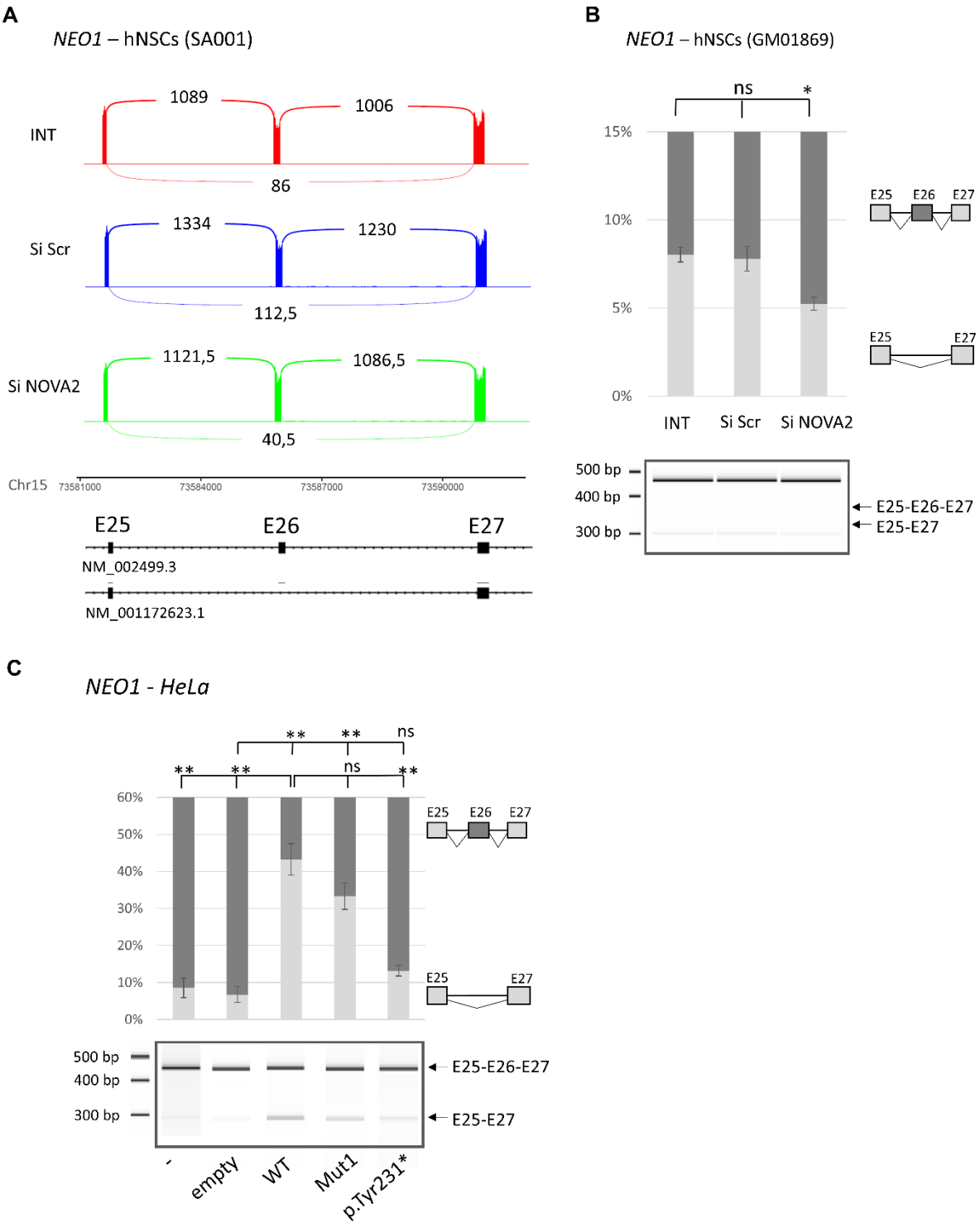

Figure S5. Additional examples of alternative splicing events regulated by NOVA2 in human cells*: NEO1* exon 26

(A) Sashimi plot established from RNA-Seq data for *NEO1* exon 26 is represented. The number of reads supporting the existence of each exon-exon junction is indicated as an average between

data from the two independent series of hNSCs treated with Interferin only (in red), with scramble siRNA (in blue) or with *NOVA2* siRNA (in green) during 48 hours. **(B)** Confirmation of the consequences of *NOVA2* inactivation on these AS events in another hNSC cell line (GM01869), transfected with siScramble (si Scr), si*NOVA2*, or treated with the transfecting agent only (INT). The RT-PCR products obtained were analyzed by migration on a 2,100 Bioanalyzer instrument (Agilent Technology). Experiments were done in triplicates. The error bars indicate the SEM. Kruskal–Wallis' ANOVA with Dunn's multiple comparison test was performed * $p<0.05$, ns: non-significant. **(C)** Effect of overexpression of wild-type or mutant NOVA2 proteins on these AS events in HeLa cells. The RT-PCR products obtained were analyzed by migration on a 2,100 Bioanalyzer instrument (Agilent Technology). Four series of experiments were analyzed. The error bars indicate the SEM. Brown-Forsythe and Welch's ANOVA with Holm-Sidak's multiple comparisons *** $p<0.001$, **$p<0.01$, * $p<0.05$, ns: non-significant.

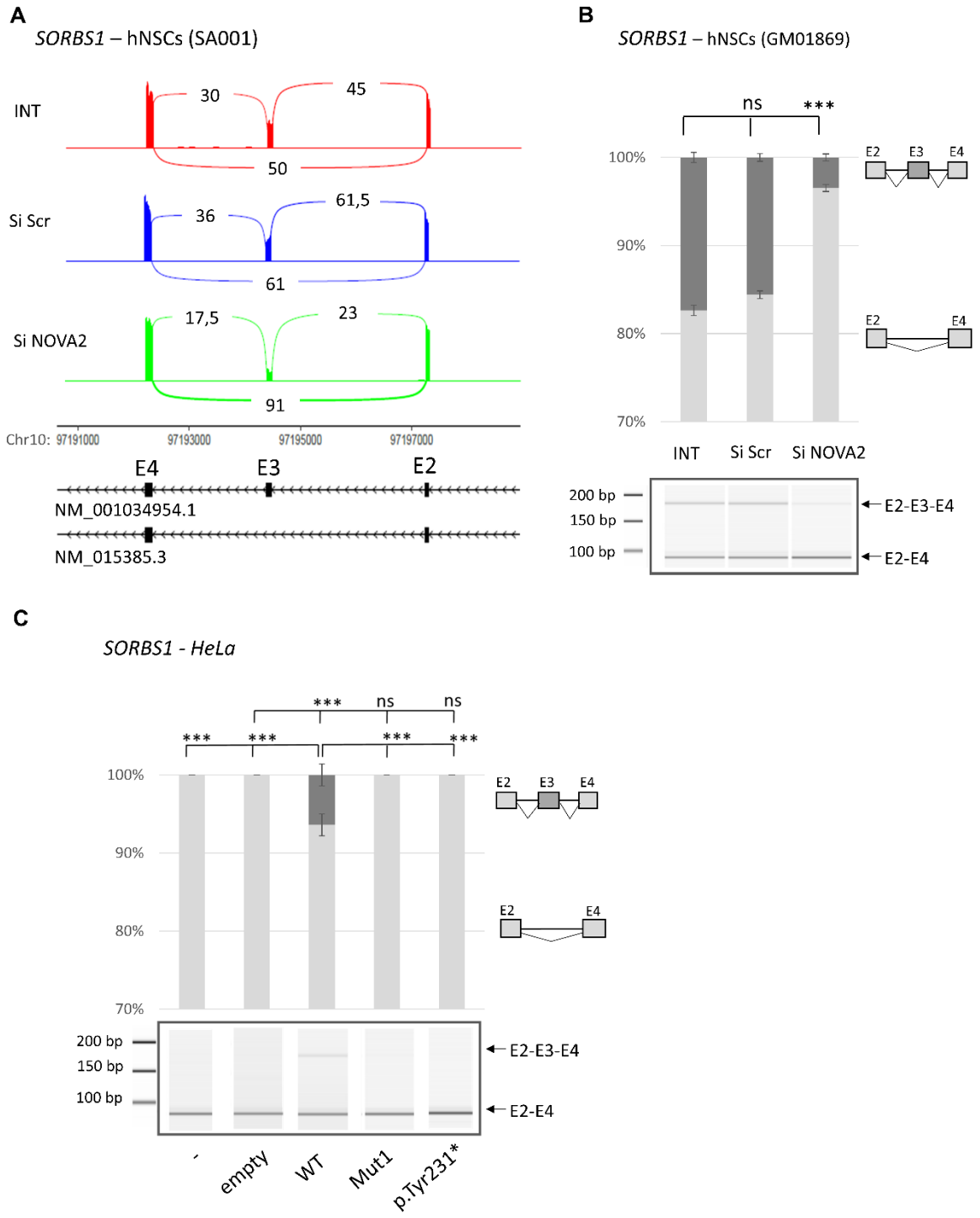

Figure S6. Additional examples of alternative splicing events regulated by NOVA2 in human cells*: SORBS1* exon 3

(A) Sashimi plot established from RNA-Seq data for *SORBS1* exon 3 is represented. The number of reads supporting the existence of each exon-exon junction is indicated as an average between data from the two independent series of hNSCs treated with Interferin only (in red),

with scramble siRNA (in blue) or with *NOVA2* siRNA (in green) during 48 hours. **(B)** Confirmation of the consequences of *NOVA2* inactivation on these AS events in another hNSC cell line (GM01869), transfected with siScramble (si Scr), si*NOVA2*, or treated with the transfecting agent only (INT). The RT-PCR products obtained were analyzed by migration on a 2,100 Bioanalyzer instrument (Agilent Technology). Experiments were done in triplicates. The error bars indicate the SEM. Kruskal–Wallis' ANOVA with Dunn's multiple comparison test was performed *** $p<0.001$, ns: non-significant **(C)** Effect of overexpression of wild-type or mutant NOVA2 proteins on these AS events in HeLa cells. The RT-PCR products obtained were analyzed by migration on a 2,100 Bioanalyzer instrument (Agilent Technology). Four series of experiments were analyzed. The error bars indicate the SEM. Brown-Forsythe and Welch's ANOVA with Holm-Sidak's multiple comparisons *** $p<0.001$, ns: non-significant.

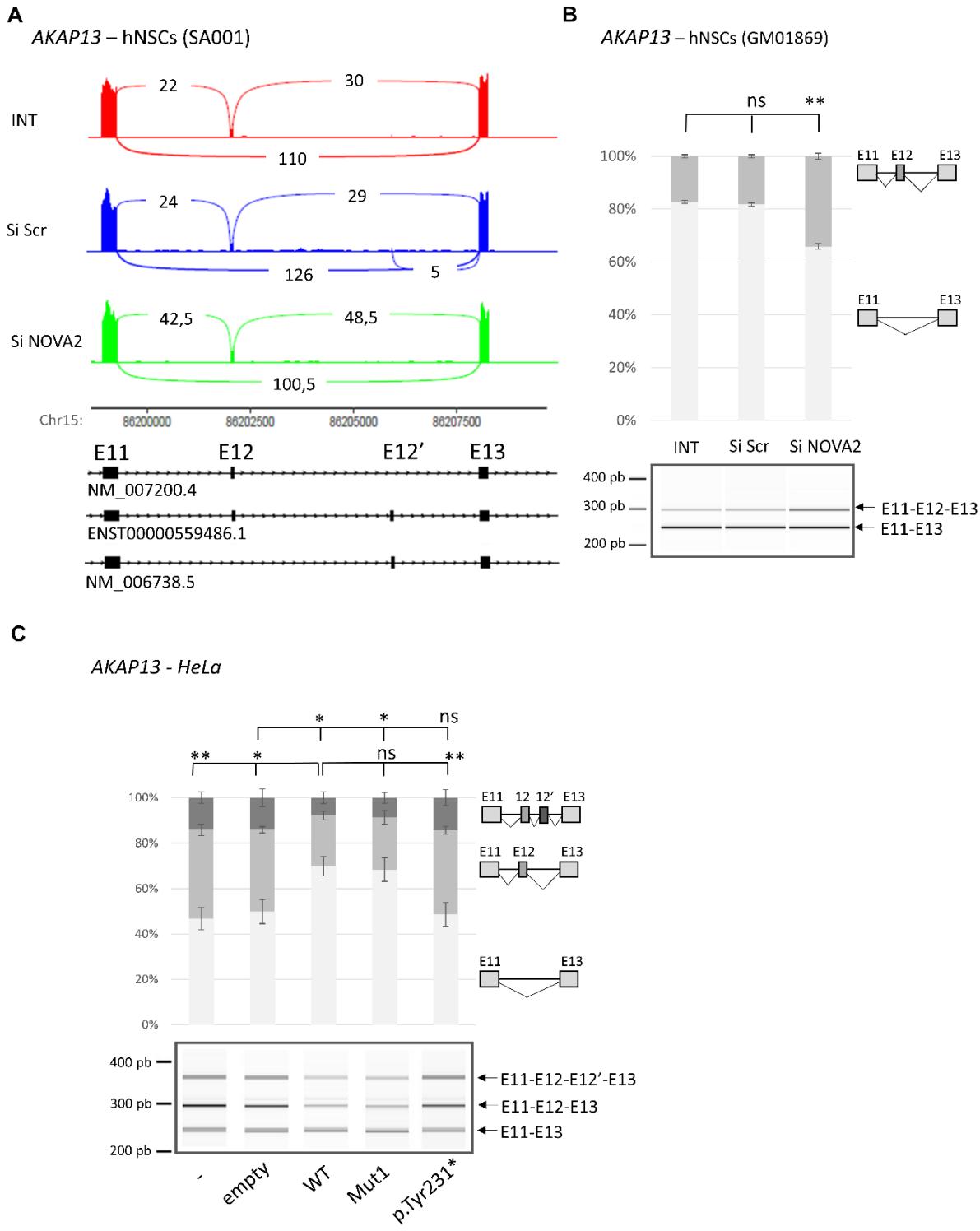

Figure S7. Additional examples of alternative splicing events regulated by NOVA2 in human cells*: AKAP13* exon 12.

(**A**) Sashimi plot established from RNA-Seq data for *AKAP13* exon 12 is represented. The number of reads supporting the existence of each exon-exon junction is indicated as an average

between data from the two independent series of hNSCs treated with Interferin only (in red), with scramble siRNA (in blue) or with *NOVA2* siRNA (in green) during 48 hours. **(B)** Confirmation of the consequences of *NOVA2* inactivation on these AS events in another hNSC cell line (GM01869), transfected with siScramble (si Scr), si*NOVA2*, or treated with the transfecting agent only (INT). The RT-PCR products obtained were analyzed by migration on a 2,100 Bioanalyzer instrument (Agilent Technology). Experiments were done in triplicates. The error bars indicate the SEM. Kruskal–Wallis' ANOVA with Dunn's multiple comparison test was performed **$p<0.01$, ns: non-significant **(C)** Effect of overexpression of wild-type or mutant NOVA2 proteins on these AS events in HeLa cells. The RT-PCR products obtained were analyzed by migration on a 2,100 Bioanalyzer instrument (Agilent Technology). Four series of experiments were analyzed. The error bars indicate the SEM. Brown-Forsythe and Welch's ANOVA with Holm-Sidak's multiple comparisons **$p<0.01$, * $p<0.05$, ns: non-significant.

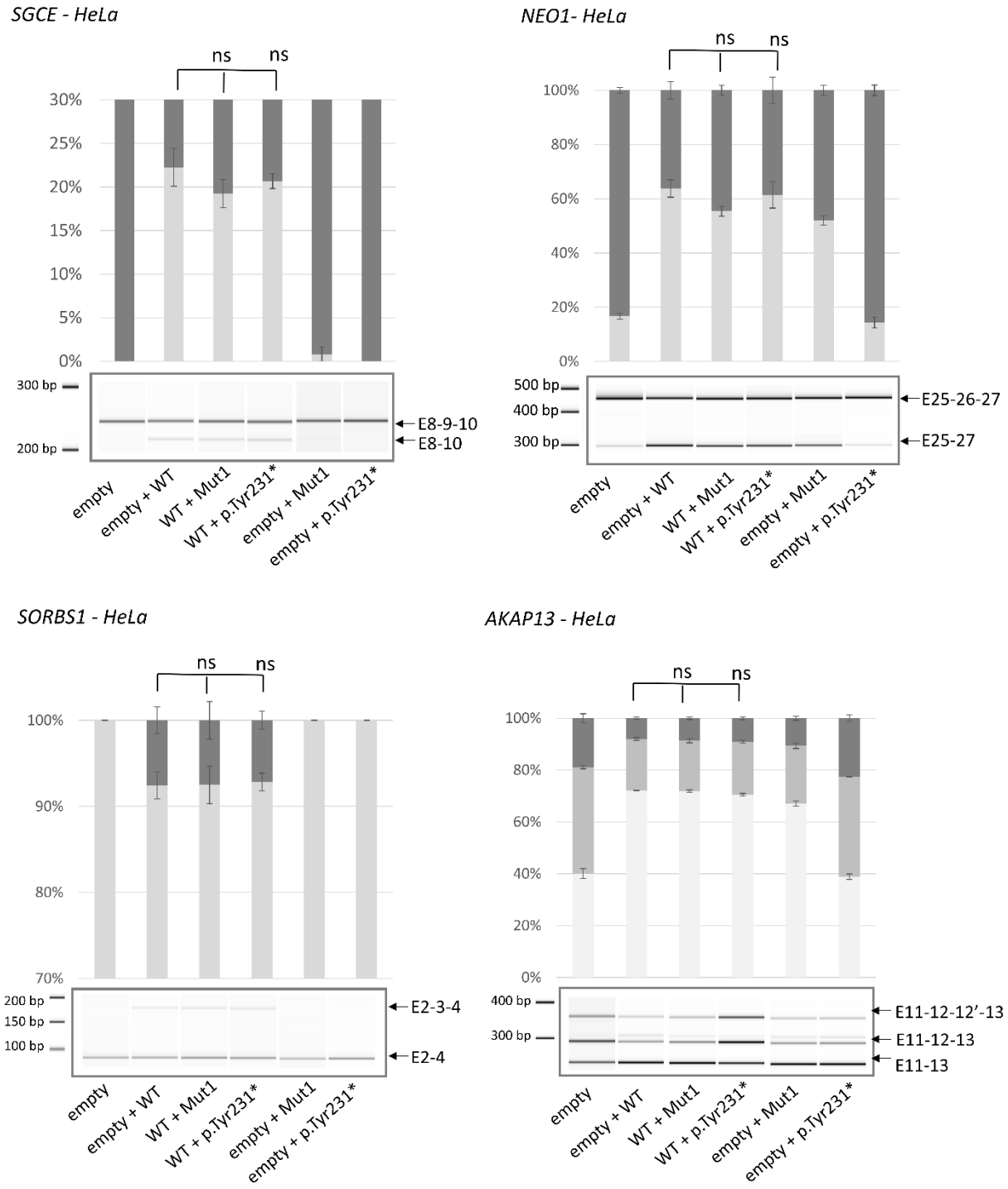

Figure S8

Figure S8. Effect of co-expression of wild-type and variant NOVA2 proteins on the regulation of AS events in HeLa cells

HeLa cells were transfected with different combination of empty plasmid (empty), EGFP-tagged NOVA2 wild-type (WT) and variant (Mut1 or p.Tyr231*) proteins to check for a potential dominant-negative effect of NOVA2 variants on the regulation of AS events in *SGCE, NEO1, SORBS1* and *AKAP13*. The RT-PCR products obtained were analyzed by migration on a 2,100 Bioanalyzer instrument (Agilent Technology). Three series of experiments were analyzed. The error bars indicate the SEM. Kruskal–Wallis' ANOVA with Dunn's multiple comparison test was performed ns: non-significant.

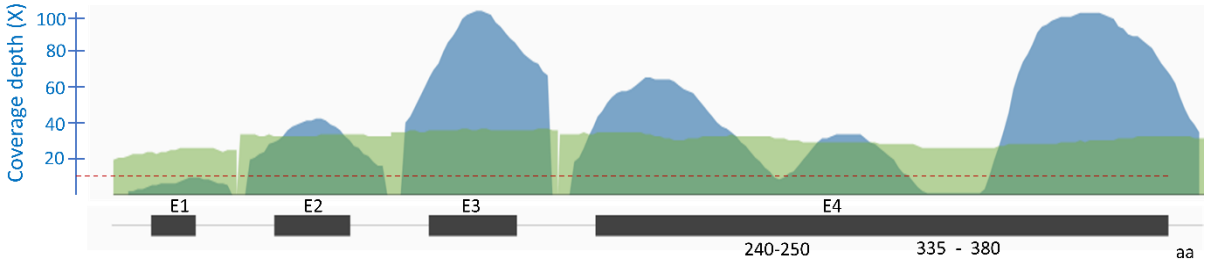

Figure S9. Depth of coverage along the *NOVA2* gene obtained by WES (extracted from gnomAD)

Representation of the depth of coverage of *NOVA2* coding sequences which can be obtained by Whole Exome Sequencing (WES, in blue) or Whole Genome Sequencing (WGS, in green)(data extracted from gnomAD: https://gnomad.broadinstitute.org/gene/ ENSG00000104967). Two regions of exon 4 are poorly (<10X, red line) or not covered: c.720 to 750, encoding amino acids 240 to 250, and c.1005 to 1140 encoding amino acids 355 to 380.