Supporting Information

Bogaert et al. 10.1073/pnas.1705895114

### Between-Groups Comparisons Using Univariate Tests

The two-tailed results of the nonparametric Mann–Whitney U test were described in the main text and the full results, including mean ranks, are shown in Table S1.

Between-Groups Comparisons Using Jonckheere–Terpstra Tests

The omnibus, two-tailed results of the nonparametric Jonckheere–Terpstra test for ordered alternatives are discussed in the main text and *Materials and Methods*. The standardized test statistic (STS) for the omnibus test was: anti-NLGN4Y isoform 1, STS = 3.90, P = 0.000096; batch-corrected anti-NLGN4Y isoform 2, STS = 2.17, P = 0.030; and combined anti-NLGN4Y, STS = 3.23, P = 0.0012. The dependent variables in these analyses were antibody concentrations for the three anti-NLGN4Y measures controlling for pregnancy in the subsample of women (n = 142). To compute these dependent variables, regression analyses were run with each of the antibody variables as the criterion and number of pregnancies as the predictor. The standardized residual was saved and was then used as the dependent variable in the Jonckheere–Terpstra tests. The P values of the follow-up tests were one-tailed, in keeping with the directional nature of the Jonckheere–Terpstra test. The results are shown in Fig. 2 (for anti-NLGN4Y isoform 1) and Table S2 (both isoforms and their combination) and are discussed in the main text.

Scatterplots of Anti-NLGN4Y Antibody Concentrations, Controlling and Not Controlling for Pregnancy

In addition to the presentation of rank orders of data in the main text (Figs. 1 and 2) and the raw values in Fig. S1, we present the scatterplots of antibody concentrations not controlling for pregnancy by group (Fig. S4 *A*, *C*, and *E*), and the scatterplots of antibody concentrations controlling for pregnancy by group (Fig. S4 *B*, *D*, and *F*). For scatterplots of antibody concentrations not controlling for pregnancy (Fig. S4 *A*, *C*, and *E*), raw levels of anti-NLGN4Y (ng/mL) are displayed for isoform 1 and isoform 2 (Fig. S4 *A* and *C*). It is important to note that anti-NLGN4Y isoform 2 is not corrected for batch effects (Fig. S4*C*). For combined anti-NLGN4Y, the (uncorrected) standardized scores (z-scores) for anti-NLGN4Y isoform 1 were added to the (batch-corrected) z-scores for isoform 2. For scatterplots of antibody concentrations controlling for pregnancy (Fig. S4 *B*, *D*, and *F*), standardized residuals were saved after conducting regression analyses with each of the antibody variables as the criterion variable and number of pregnancies as the predictor. The differences between groups are difficult to see in these scatterplots because of the effect of the extreme cases on the scaling of the y axis. Thus, we have added a dashed reference line below the lowest observed value, making it easier to see that the groups generally differ at the lower ends of their distributions, particularly when pregnancy is controlled for (Fig. S4 *B*, *D*, and *F*).
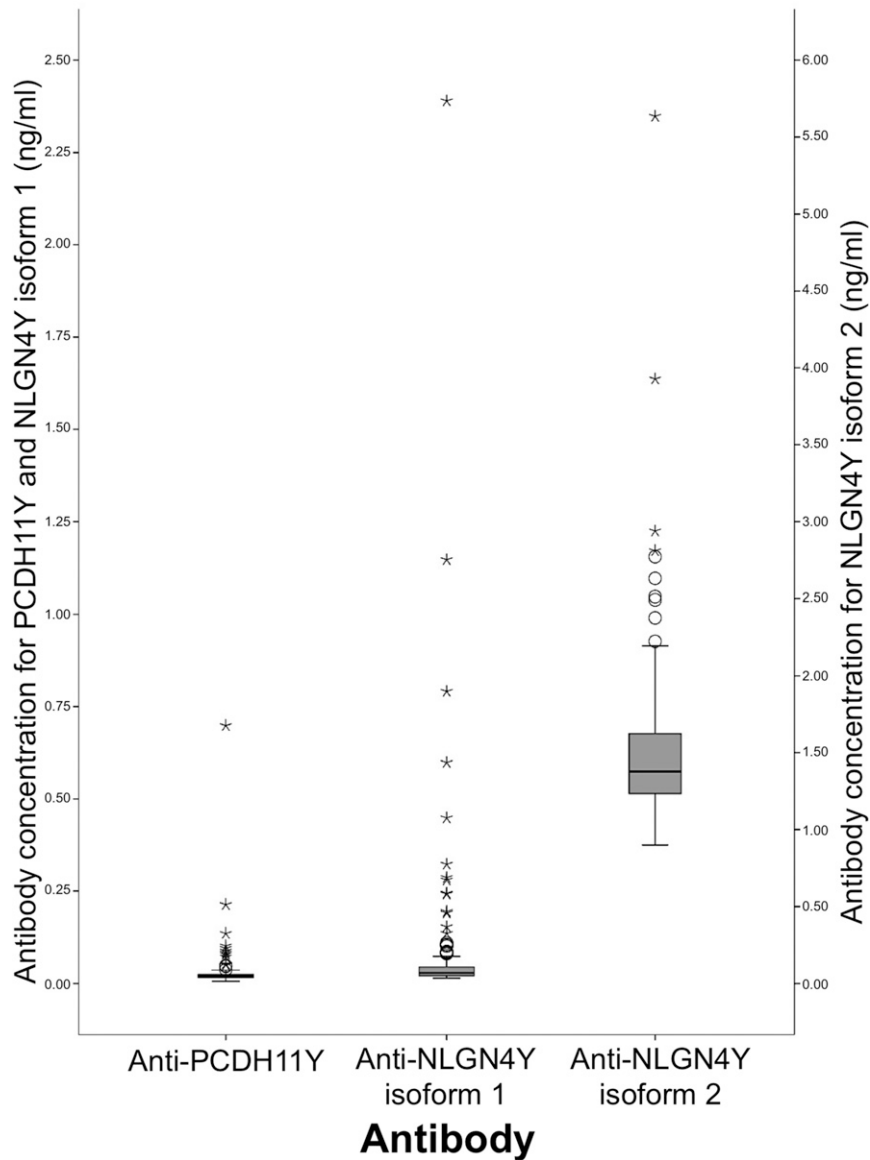
Fig. S1. Boxplot of raw levels of antibody concentrations (ng/mL) to each protein. Raw levels of anti-PCDH11Y (mean = 0.03 ng/mL, SD = 0.06 ng/mL, median = 0.02 ng/mL), anti-NLGN4Y isoform 1 (mean = 0.08 ng/mL, SD = 0.23 ng/mL, median = 0.03 ng/mL), and anti-NLGN4Y isoform 2 (mean = 1.53 ng/mL, SD = 0.54 ng/mL, median = 1.38 ng/mL) across study groups (n = 154). Levels of anti-NLGN4Y isoform 2 are not batch-corrected. Note the dual axes. The box represents the interquartile range (IQR), where the middle 50% of scores lie (26). The top of the box is the upper quartile (UQ), and the bottom of the box is the lower quartile (LQ). The line inside the box is the median. A circle (O) represents an outlier [i.e., score > UQ + (IQR × 1.5)]. An asterisk (*) represents an extreme case [i.e., score > UQ + (IQR × 3)]. The "whiskers" are the lines extending from the top and bottom of the box, and they represent approximately the top 25% and the bottom 25% of where the data points fall, excluding the outliers and extreme cases.

```
NLGN4Y_isoform_2_256aa    MLRPQGLLWLPLLFTSVCVMLNSNVLLWITALAIKFTLIDSQAQYPVVNTNYGKIQGLRT
NLGN4Y_isoform_1_816aa    MLRPQGLLWLPLLFTSVCVMLNSNVLLWITALAIKFTLIDSQAQYPVVNTNYGKIQGLRT
                          ************************************************************

NLGN4Y_isoform_2_256aa    PLPSEILGPVEQYLGVPYASPPTGERRFQPPESPSSWTGIRNATQFSAVCPQHLDERFLL
NLGN4Y_isoform_1_816aa    PLPSEILGPVEQYLGVPYASPPTGERRFQPPESPSSWTGIRNATQFSAVCPQHLDERFLL
                          ************************************************************

NLGN4Y_isoform_2_256aa    HDMLPIWFTTSLDTLMTYVQDQNEDCLYLNIYVPMEDGTNIKRNADDITSNDHGEDKDIH
NLGN4Y_isoform_1_816aa    HDMLPIWFTTSLDTLMTYVQDQNEDCLYLNIYVPMED-------------------DIH
                          *************************************                   ***

NLGN4Y_isoform_2_256aa    EQNSKKPVMVYIHGGSYMEGTGNMIDGSILASYGNVIVITINYRLGILGMQEARLCGSSK
NLGN4Y_isoform_1_816aa    EQNSKKPVMVYIHGGSYMEGTGNMIDGSILASYGNVIVITINYRLGILGFLSTGDQAAKG
                          *************************************************: .:    .:.

NLGN4Y_isoform_2_256aa    MFNYFKSPFTNLINFF--------------------------------------------
NLGN4Y_isoform_1_816aa    NYGLLDQI--QALRWIEENVGAFGGDPKRVTIFGSGAGASCVSLLTLSHYSEGLFQKAII
                           :   :..    :  :.:

NLGN4Y_isoform_2_256aa    ------------------------------------------------------------
NLGN4Y_isoform_1_816aa    QSGTALSSWAVNYQPAKYTRILADKVGCNMLDTTDMVECLKNKNYKELIQQTITPATYHI


NLGN4Y_isoform_2_256aa    ------------------------------------------------------------
NLGN4Y_isoform_1_816aa    AFGPVIDGDVIPDDPQILMEQGEFLNYDIMLGVNQGEGLKFVDGIVDNEDGVTPNDFDFS


NLGN4Y_isoform_2_256aa    ------------------------------------------------------------
NLGN4Y_isoform_1_816aa    VSNFVDNLYGYPEGKDTLRETIKFMYTDWADKENPETRRKTLVALFTDHQWVAPAVATAD


NLGN4Y_isoform_2_256aa    ------------------------------------------------------------
NLGN4Y_isoform_1_816aa    LHAQYGSPTYFYAFYHHCQSEMKPSWADSAHGDEVPYVFGIPMIGPTELFSCNFSKNDVM


NLGN4Y_isoform_2_256aa    ------------------------------------------------------------
NLGN4Y_isoform_1_816aa    LSAVVMTYWTNFAKTGDPNQPVPQDTKFIHTKPNRFEEVAWSKYNPKDQLYLHIGLKPRV


NLGN4Y_isoform_2_256aa    ------------------------------------------------------------
NLGN4Y_isoform_1_816aa    RDHYRATKVAFWLELVPHLHNLNEIFQYVSTTTKVPPPDMTSFPYGTRRSPAKIWPTTKR


NLGN4Y_isoform_2_256aa    ------------------------------------------------------------
NLGN4Y_isoform_1_816aa    PAITPANNPKHSKDPHKTGPEDTTVLIETKRDYSTELSVTIAVGASLLFLNILAFAALYY



NLGN4Y_isoform_2_256aa    ------------------------------------------------------------
NLGN4Y_isoform_1_816aa    KKDKRRHETHRHPSPQRNTTNDITHIQNEEIMSLQMKQLEHDHECESLQAHDTLRLTCPP


NLGN4Y_isoform_2_256aa    -----------------------------------------------------------
NLGN4Y_isoform_1_816aa    DYTLTLRRSPDDIPFMTPNTITMIPNTLMGMQPLHTFKTFSGGQNSTNLPHGHSTTRV
```
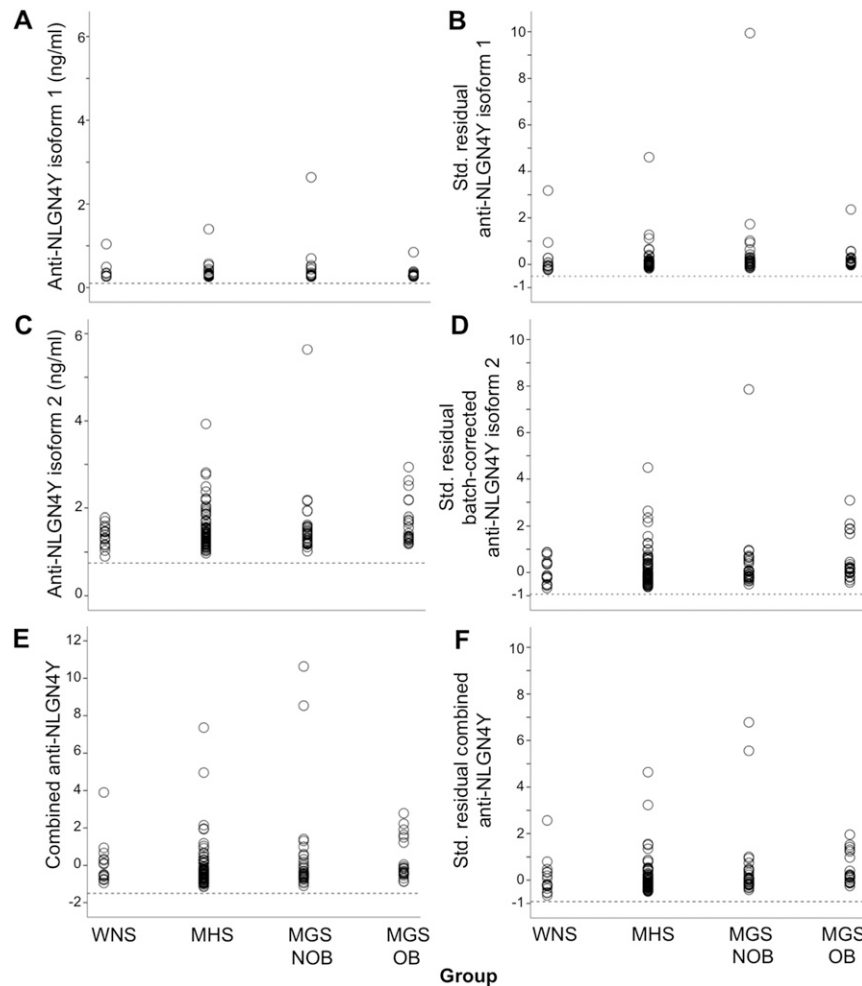
Fig. S2. Alignment of NLGN4Y isoform protein sequences. Amino acid alignment (Clustal Omega) of full NLGN4Y isoform protein sequences (isoform 1 length = 816 aa; isoform 2 length = 256 aa; source: National Center for Biotechnology Information). Full sequences for each isoform are shown, and fragments produced for the development of recombinant proteins employed in ELISA are colored for each. Blue = isoform 1; red = isoform 2; an asterisk (*) = identical; a colon (:) = conserved substitution; and a period (.) = semiconserved substitution.

Fig. S3. Standard curves for each assay. (*A*) Standard curve for PCDH11Y across all batches. (*B*) Standard curve for NLGN4Y isoform 1 across all batches. (*C*) Standard curve for NLGN4Y isoform 2 across all batches.

Fig. S4. Scatterplots of antibody concentrations not controlling (*A*, *C*, and *E*) and controlling (*B*, *D*, and *F*) for pregnancy for NLGN4Y by group (*n* = 142). MGS NOB, mothers of gay sons with no older brothers (*n* = 31); MGS OB, mothers of gay sons with older brothers (*n* = 23); MHS, mothers of heterosexual sons (*n* = 72); Std. residual, Standardized residual; WNS, women with no sons (*n* = 16). The dashed reference line runs below the lowest observed value across groups. (*A*) Raw levels of anti-NLGN4Y isoform 1 (ng/mL) by group. (*B*) Standardized residual controlling for pregnancy for NLGN4Y isoform 1 by group. (*C*) Raw levels of anti-NLGN4Y isoform 2 (ng/mL) by group. Levels of anti-NLGN4Y isoform 2 are not batch-corrected. (*D*) Standardized residual controlling for pregnancy for NLGN4Y isoform 2 by group and corrected for batch effects. (*E*) Standardized scores of combined anti-NLGN4Y by group. (*F*) Standardized residual controlling for pregnancy for combined NLGN4Y by group.

Table S1.   Results of Mann–Whitney *U* test on antibody levels with sex of participant as the independent (i.e., grouping) variable (*n* = 154)

Antibody	Mann–Whitney *U*	*Z*	*P**	Mean rank: men	Mean rank: women
Anti-PCDH11Y	824.50	−0.19	0.855	75.21	77.69
Anti-NLGN4Y isoform 1	461.00	−2.64	0.007	44.92	80.25
Batch-corrected anti-NLGN4Y isoform 2	460.00	−2.64	0.007	44.83	80.26
Combined anti-NLGN4Y	420.00	−2.91	0.003	41.50	80.54

A negative value for the standardized ("*Z*") test statistic indicates that the mean rank for women was higher than the mean rank for men.
*Two-tailed *P* values are indicated.

Table S2. Pairwise comparisons of antibody levels with significant Jonckheere-Terpstra test results controlling for pregnancy (n = 142)

Group 1	Group 2	Anti-NLGN4Y isoform 1		Batch-corrected anti-NLGN4Y isoform 2		Combined anti-NLGN4Y	
		Standardized test statistic	P	Standardized test statistic	P	Standardized test statistic	P
Women with no sons	Mothers of heterosexual sons	1.53	0.064	0.15	0.440	0.39	0.348
Women with no sons	Mothers of gay sons with no older brothers	2.04	0.021	0.65	0.258	1.19	0.117
Women with no sons	Mothers of gay sons with older brothers	2.43	0.008	1.34	0.090	2.00	0.023
Mothers of heterosexual sons	Mothers of gay sons with no older brothers	1.97	0.024	0.86	0.196	1.70	0.045
Mothers of heterosexual sons	Mothers of gay sons with older brothers	3.39	0.00035	2.29	0.011	3.08	0.001
Mothers of gay sons with no older brothers	Mothers of gay sons with older brothers	0.69	0.245	1.65	0.049	1.67	0.047

A positive value for the standardized test statistic for pairwise comparisons indicates that the mean rank for group 2 was higher than the mean rank for group 1. A negative value indicates the opposite. One-tailed P values are indicated.

Table S3. Age at examination for participants assigned to study groups (n = 154)

Group	N	Mean	SD
Men*	12	26.36	11.03
Women with no sons	16	39.63	13.79
Mothers of heterosexual sons	72	49.42	11.09
Mothers of gay sons with no older brothers	31	54.16	10.09
Mothers of gay sons with older brothers	23	53.35	9.49

*One male subject neglected to report his age, so the mean age (and SD) for men were computed on 11 cases.

Table S4. Spearman's correlations of all pregnancies with ranked antibody measures within all women (n = 142) and within mothers only (n = 137)

Antibody measure	All women (n = 142)		Mothers only (n = 137)	
	ρ	P	ρ	P
Ranked anti-PCDH11Y	0.024	0.781	0.005	0.958
Ranked anti-NLGN4Y isoform 1	−0.153	0.068	−0.200	0.019
Ranked batch-corrected anti-NLGN4Y isoform 2	−0.164	0.051	−0.209	0.014
Ranked combined anti-NLGN4Y	−0.176	0.036	−0.229	0.007

Table S5. Total number of pregnancies within women assigned to study groups (n = 142)

Group	n	Mean	SD
Women with no sons	16	1.63	1.31
Mothers of heterosexual sons*	72	2.72	1.24
Mothers of gay sons with no older brothers*	31	3.10	1.74
Mothers of gay sons with older brothers*,†	23	3.78	1.91

*Significantly different from "women with no sons."
†Significantly different from "mothers of heterosexual sons" in a one-way ANOVA (see *Materials and Methods*).