Supplemental Materials

Molecular Biology of the Cell

Major et al.

Combined Supplemental Materials

Supplemental Materials References

Horton, P., and Nakai, K. (1997). Better prediction of protein cellular localization sites with the k nearest neighbors classifier. Proc Int Conf Intell Syst Mol Biol 5, 147-152.

Kosugi, S., Hasebe, M., Tomita, M., and Yanagawa, H. (2009). Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. Proc Natl Acad Sci U S A 106, 10171-10176.

Major, A.T., Whiley, P.A., and Loveland, K.L. (2011). Expression of nucleocytoplasmic transport machinery: clues to regulation of spermatogenic development. Biochim Biophys Acta 1813, 1668-1688.

Nakai, K., and Horton, P. (1999). PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem Sci 24, 34-36.

Namekawa, S.H., Park, P.J., Zhang, L.F., Shima, J.E., McCarrey, J.R., Griswold, M.D., and Lee, J.T. (2006). Postmeiotic sex chromatin in the male germline of mice. Curr Biol 16, 660-667.

Reinhardt, A., and Hubbard, T. (1998). Using neural networks for prediction of the subcellular location of proteins. Nucleic Acids Res 26, 2230-2236.

Shima, J.E., McLean, D.J., McCarrey, J.R., and Griswold, M.D. (2004). The murine testicular transcriptome: characterizing gene expression in the testis during the progression of spermatogenesis. Biol Reprod 71, 319-330.

TABLE S1: Outcomes of cNLS Mapper and PSORT II predictions for candidate IMPα2 cargo proteins.

Gene symbol Gene name	cNLS mapper		PSORT II Prediction			
	Mono-partite	Bi-partite	Reinhart's Method	Reliability	k-NN Prediction (Primary)	k-NN Prediction (Others)
PSPC1 paraspeckle protein 1	ND	3.8	nuclear	76.7	73.9 %: nuclear	8.7 %: cytoplasmic 4.3 %: vesicles of secretory system 4.3 %: Golgi 4.3 %: cytoskeletal 4.3 %: mitochondrial
RRP1 ribosomal RNA processing 1 homolog	18	15.2	nuclear	89	91.3 %: nuclear	4.3 %: vesicles of secretory system 4.3 %: cytoskeletal
MAGED-1 Maged1 melanoma antigen, family D, 1	2.5	6.3	nuclear	89	69.6 %: nuclear	13.0 %: plasma membrane 8.7 %: cytoplasmic 4.3 %: vesicles of secretory system 4.3 %: Golgi
HNRPC heterogeneous nuclear ribonucleoprotein C	5.5	5	nuclear	94.1	47.8 %: nuclear	34.8 %: mitochondrial 8.7 %: cytoplasmic 4.3 %: vesicles of secretory system 4.3 %: cytoskeletal
BOD1L biorientation of chromosomes in cell division 1-like	12	9.7	nuclear	94.1	82.6 %: nuclear	8.7 %: plasma membrane 4.3 %: cytoplasmic 4.3 %: cytoskeletal

The putative monopartite or bipartite NLS scores were determined using cNLS Mapper (Kosugi et al., 2009). ND: Samples with a score of 2.0 or below (no detectable NLS). The PSORT II (Nakai and Horton, 1999) data includes predicted localisation with reliability scores using Reinhardt's method (Reinhardt and Hubbard, 1998) and the predicted cellular compartment percentages calculated using the k Nearest Neighbors Classifier (k-NN) algorithm (Horton and Nakai, 1997).

TABLE S2: Outcomes of modulating IMPα2 expression and transport function on PSF-positive nuclear speckles.

	IMPα2-FL		IMPα2ΔIBB		IMPα2-ED		Totals
Number of cells analysed	2050		848		676		3574
Number of PSF foci +ve cells	1692		444		405		2541
Number of PSF foci detected	7424		1845		1935		11204
% PSF foci +ve cells (95% CI)	82.5% *(73.1↔91.9)%*		52.4% *(45.3↔59.4)%*		59.9% *(50.7↔69.1)%*		Lg Reg Sig (Vs ED)
Odds ratio (PSF foci +ve cells) (95% CI)	3.163 *(2.611↔3.830)*		0.735 *(0.599↔0.902)*		1.000 *(Control)*		0.000 (FL)* 0.003 (ΔIBB)*
Mean per cell values (PSF foci positive cells)	Geometric Mean (GM) & 95% CI	Ratio of GM & 95% CI	Geometric Mean (GM) & 95% CI	Ratio of GM & 95% CI	Geometric Mean (GM) & 95% CI	Ratio of GM & 95% CI	Ln Reg Sig (Vs ED)
Number of PSF foci	3.55 *(3.43↔3.67)*	0.967 *(0.896↔1.042)*	3.34 *(3.13↔3.56)*	0.909 *(0.827↔0.998)*	3.67 *(3.43↔3.93)*	1.000 *(Control)*	0.379 (FL) 0.045 (ΔIBB)
PSF foci volume sum	0.58 *(0.55↔0.61)*	0.946 *(0.854↔1.049)*	0.57 *(0.52↔0.62)*	0.923 *(0.812↔1.049)*	0.61 *(0.56↔0.67)*	1.000 *(Control)*	0.293 (FL) 0.221 (ΔIBB)
Foci PSF intensity sum	54.9 *(52.1↔57.8)*	0.962 *(0.855↔1.081)*	54.9 *(49.6↔60.7)*	0.962 *(0.832↔1.112)*	57.1 *(51.4↔63.4)*	1.000 *(Control)*	0.512 (FL) 0.599 (ΔIBB)
Mean values per PSF foci	Geometric Mean (GM) & 95% CI	Ratio of GM & 95% CI	Geometric Mean (GM) & 95% CI	Ratio of GM & 95% CI	Geometric Mean (GM) & 95% CI	Ratio of GM & 95% CI	GEE Sig (vs ED)
Volume	0.141 *(0.138↔0.143)*	0.975 *(0.936↔1.016)*	0.146 *(0.140↔0.152)*	1.014 *(0.960↔1.070)*	0.144 *(0.139↔0.150)*	1.000 *(Control)*	0.236 (FL) 0.621 (ΔIBB)
PSF intensity sum	569 *(556↔583)*	0.966 *(0.909↔1.025)*	612 *(580↔646)*	1.037 *(0.961↔1.120)*	590 *(558↔623)*	1.000 *(Control)*	0.252 (FL) 0.347 (ΔIBB)
PSF voxel intensity	97.2 *(96.2↔98.2)*	0.992 *(0.964↔1.021)*	100.2 *(97.7↔102.8)*	1.022 *(0.986↔1.060)*	98.0 *(95.4↔100.7)*	1.000 *(Control)*	0.575 (FL) 0.237 (ΔIBB)
Trend relative to IMPα2-ED:	**Increased foci +VE Cells (↑↑↑)**		**Decreased foci +VE Cells Cells (−↓−)**		**Normalising Control (−−−)**		

The analysed cell numbers for each group, number of detected PSF-positive nuclear foci and proportion of cells determined to contain PSF nuclear foci are presented. Values are normalised against the IMPα2 ED control (Values set to 1), an "odds ratio" identifies the fold difference in number of PSF nuclear foci positive cells compared to the IMPα2 ED sample. Data assessed on a per cell or PSF nuclear foci basis include geometric means with 95% Wald confidence intervals (95% CI), and the ratio of geometric means (GM). To determine significant differences between groups, a logistic regression (Lg Reg) model was used for PSF foci positive/ negative cells, linear regression (Ln Reg) models were used for per cell data (PSF foci positive cells) and generalised estimating equations (GEE) were used for per PSF nuclear foci data. Significance values compared to the IMPα2 ED control are shown. Using Bonferoni correction, the significance threshold was reassigned from 0.05 to 0.016 (0.05 ÷ 3 transfection groups). Using these criteria, the only significantly different values observed are changes in the percentage of cells positive for PSF nuclear bodies (*).

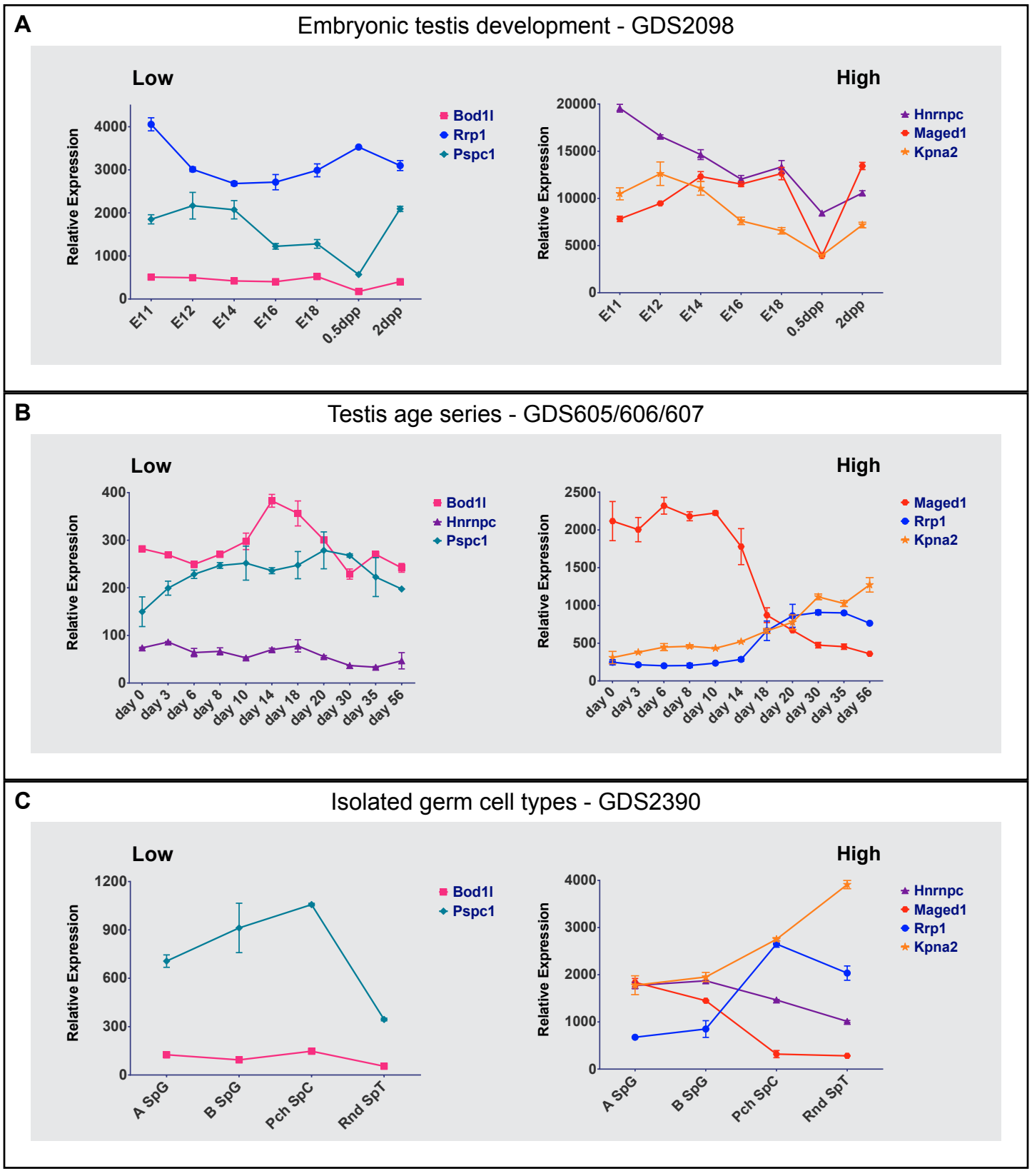

FIGURE S1: Expression of the putative IMPα2 cargoes throughout embryonic testis development and spermatogenesis. Single probesets representing each candidate IMPα2 cargo gene were selected from each GEO dataset in a systematic manner as previously described (Major et al., 2011). Data are separated into lower/higher detection levels for visualisation and error bars correspond to standard error of the mean (SEM), where multiple values were present. (**A**) Embryonic testis development (GDS2098), generated by Gaido and Lehmann (no linked publication) from normal whole mouse testes embryonic day (E) 11 through to 2 days post partum (dpp). Probesets shown here are as follows: Kpna2 (1415860_at), Pspc1 (1423192_at), Rrp1 (1427720_a_at), Maged1 (1450062_a_at), Hnrnpc (1460240_a_at) and Bod1l (1460736_at). (**B**) Mouse testis age series are from the GEO datasets GDS605, GDS606 and GDS607 (Shima et al., 2004). Probesets shown here are as follows: Kpna2 (GDS605 92790_at), Pspc1 (GDS605 103393_at), Rrp1 (GDS605 98604_at), Maged1 (GDS605 96703_at), Hnrnpc (GDS605 160199_at) and Bod1l (GDS606 113843_at). (**C**) Isolated mouse germ cell types from GEO dataset GDS2390, containing Type A spermatogonia (A SpG), Type B spermatogonia (B SpG), pachytene spermatocytes (Pch Spc) and round spermatids (Rnd SpT) (Namekawa et al., 2006). Probesets shown here are as follows: Kpna2 (1415860_at), Pspc1(1423192_at), Rrp1 (1427720_a_at), Maged1 (1450062_a_at), Hnrnpc (1460240_a_at) and Bod1l (1460736_at).

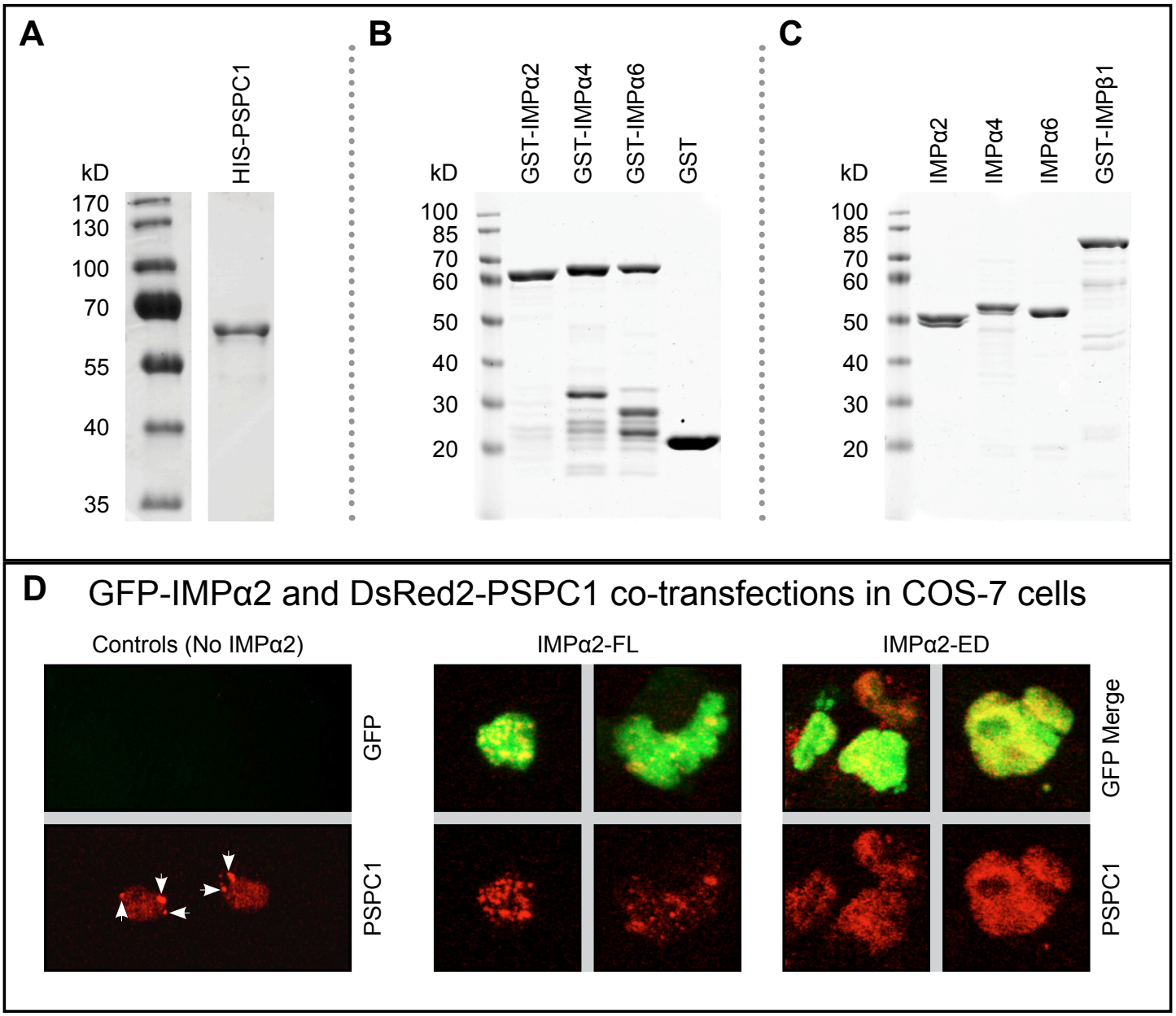

FIGURE S2: Purified recombinant tagged proteins and COS-7 cell transient co-transfections. Bacterially expressed, affinity purified, recombinant proteins used for ELISA-based importin binding assays, separated by SDS-PAGE and stained with Coomassie blue (**A-C**). Size markers are indicated to the left of each gel in kilodaltons (kD). (**A**) HIS-tagged PSPC1 protein. (**B**) GST alone or GST-tagged IMPα proteins. (**C**) GST-cleaved IMPα proteins and the GST-tagged IMPβ1 protein to which they were pre-hybridised for the ELISA based binding assay. (**D**) COS 7 cells were transiently transfected with plasmids encoding DsRed2-PSPC1 (PSPC1) plus GFP-IMPα2 full length (IMPα2-FL), mutant (IMPα2-ED) constructs or no IMPα2 plasmid. Images were acquired via confocal laser scanning microscopy (CLSM) two days post transfection, with two examples shown for each experimental group. Several paraspeckles are marked with white arrows. At least 20 COS-7 cells in each transfection group were scored for either the presence or absence of paraspeckles.

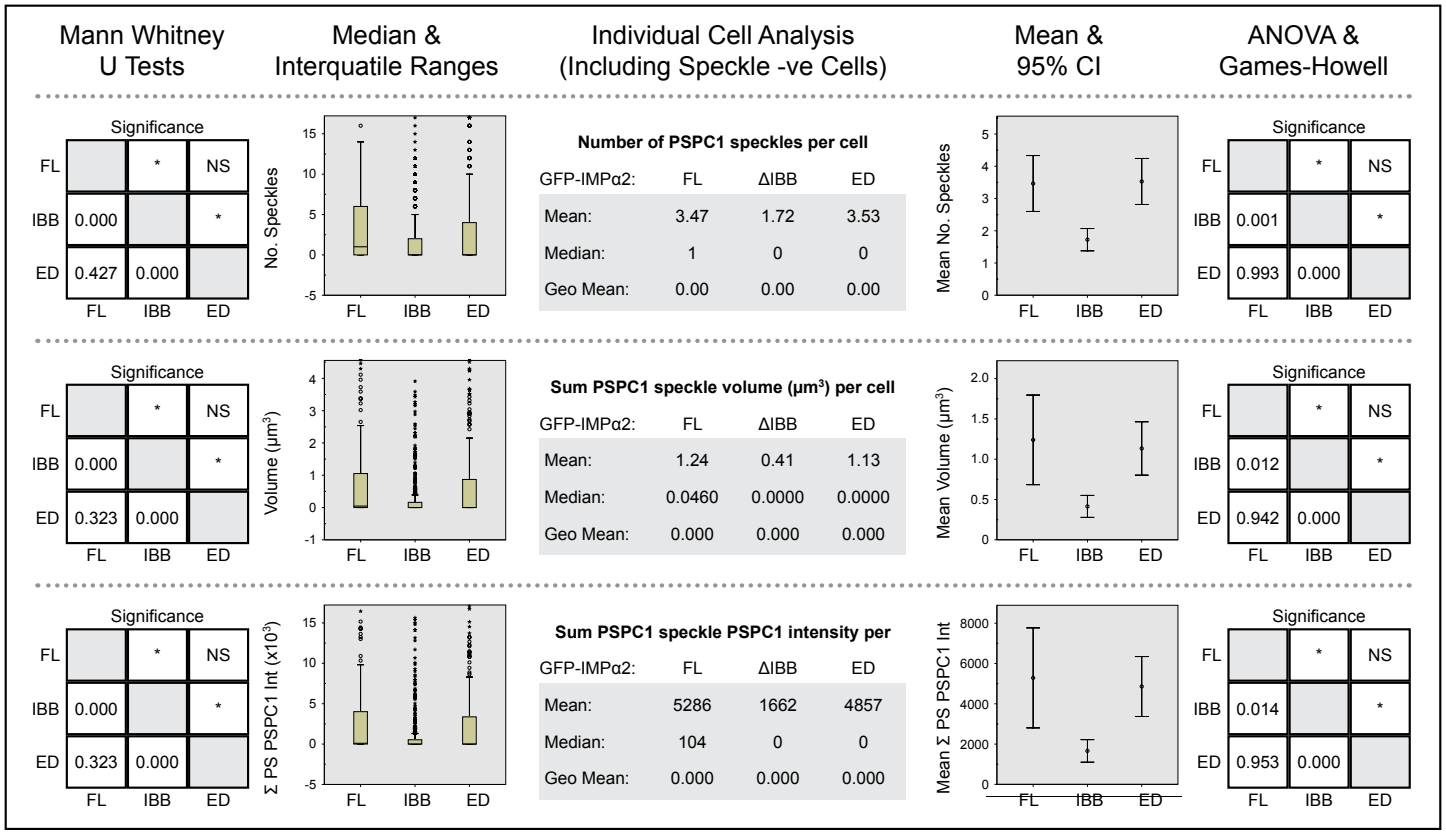

FIGURE S3: IMPα2 activity modulates PSPC1 nuclear speckle number/size – statistics based on all cells including PSPC1 speckle negative cells. Detailed results of PSPC1 speckle assessment and quantification performed for each transfection group analysed (GFP-IMPα2-FL, GFP-IMPα2ΔIBB or GFP-IMPα2-ED). Data were obtained by performing image analysis of the PSPC1 immunofluorescence signal using Imaris software to perform a spots analysis. Each measure is displayed in several formats: (1) boxplots showing the median and interquatile ranges, the whiskers extend to cover theoretical 95% of data if assuming normal distribution with outliers shown as points and extreme outliers shown as asterisks, (2) Median, Mean and Geometric (Geo) mean values for each group, (3) Plots showing arithmetic mean value & 95% confidence intervals (CI) for each transfection group, (4) Statistical significance between groups, as determined by parametric (One-Way ANOVA with a Games-Howell posthoc test) and non-parametric tests (Mann Whitney U). Using bonforeni correction the significance threshold was reassigned from 0.05 to 0.016 (0.05 ÷ 3 transfection groups), with comparisons that produced p<0.016 indicated (*). PSPC1 nuclear speckle statistics assessed on a per cell basis (including the speckle negative cells) for the number of speckles detected per cell, the sum volume of all speckles detected within any given cell and the sum of PSPC1 immunofluorescence intensities for all voxels (three dimensional pixels) of the cell determined to be part of a PSPC1 nuclear speckle.

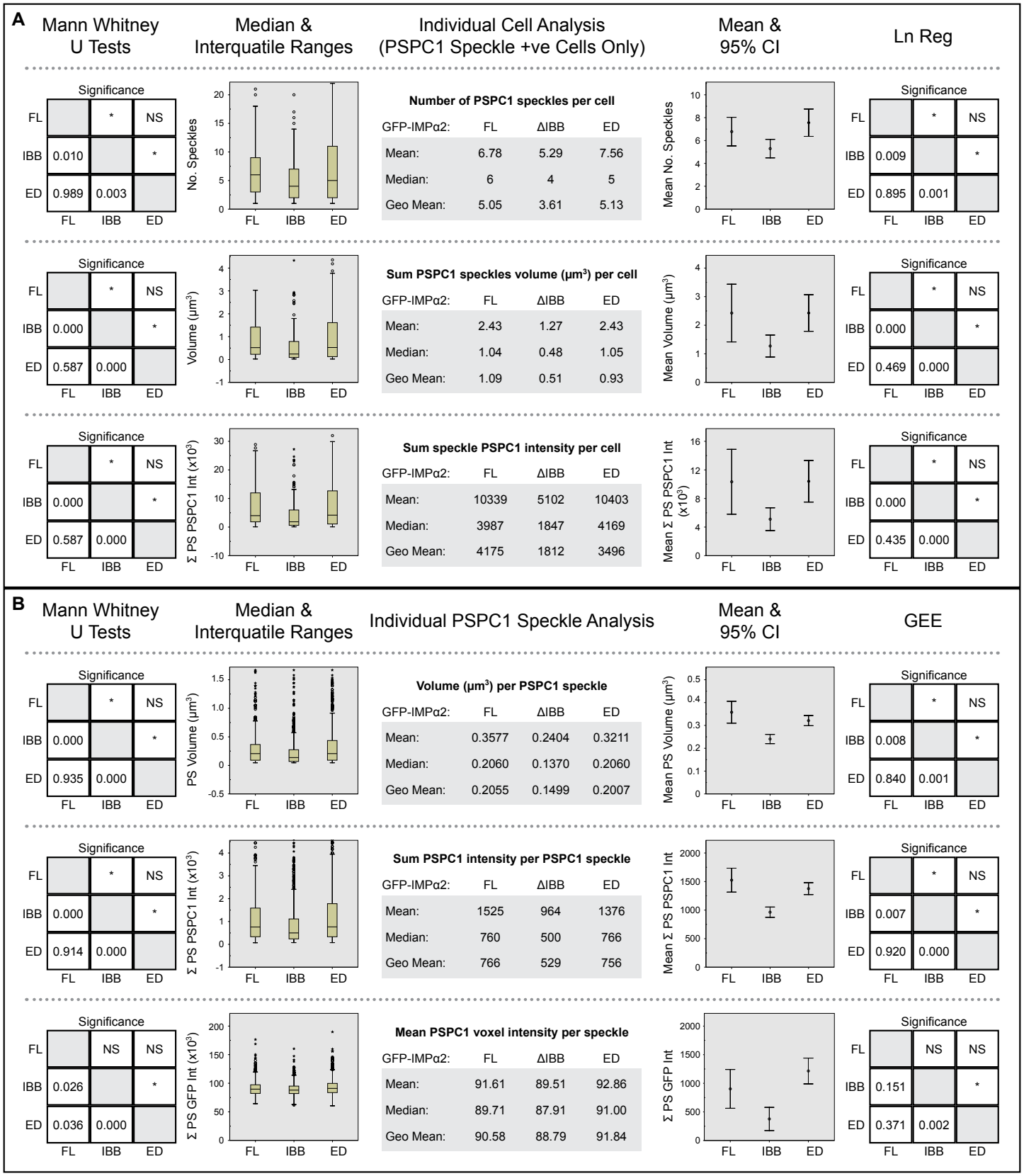

FIGURE S4: IMPα2 activity modulates PSPC1-positive nuclear speckle number/size. Detailed results of PSPC1 speckle assessment and quantification performed for each transfection group analysed (GFP-IMPα2-FL, GFP-IMPα2ΔIBB or GFP-IMPα2-ED). Data were obtained by performing image analysis of the PSPC1 immunofluorescence signal using Imaris software to perform a spots analysis. Each measure is displayed in several formats: (1) boxplots showing the median and interquatile ranges, the whiskers extend to cover theoretical 95% of data if assuming normal distribution with outliers shown as points and extreme outliers shown as asterisks, (2) Median, Mean and Geometric (Geo) mean values for each group, (3) Plots showing arithmetic mean value & 95% confidence intervals (CI) for each transfection group, (4) Statistical significance between groups was determined in multiple ways, firstly all groups were compared as determined by simple, non-parametric tests (Mann Whitney U). We also used more appropriate and sophisticated tests tailored to each group as described below. In all cases bonforeni correction was used to reassign the significance threshold from 0.05 to 0.016 (0.05 ÷ 3 transfection groups), with comparisons that produced p<0.016 indicated (*). (**A**) PSPC1 nuclear speckle statistics assessed from speckle positive cells for individual cell data, including the number of PSPC1 speckles detected per cell, the sum volume of all speckles detected within any given cell, the sum of PSPC1 immunofluorescence intensities for all voxels (three dimensional pixels) of the cell determined to be part of a speckle. Linear regression (Ln Reg) models were selected as the most appropriate way to assess statistical differences between groups. (**B**) PSPC1 nuclear speckle statistics assessed for individual PSPC1 speckles, including the volume of each speckle, the sum of PSPC1 immunofluorescence intensities for all voxels of the identified speckle and the Mean PSPC1 immunofluorescence intensity for each identified speckle. Generalised estimating equations (GEE) were selected as the most appropriate statistical assessment for differences between groups as it compensates for correlation between speckles originating from the same cell.