

Supplementary Information

Characterisation of a nucleo-adhesome

A. Byron[✉] et al.

Contents

Supplementary Figure 1

Supplementary Figure 2

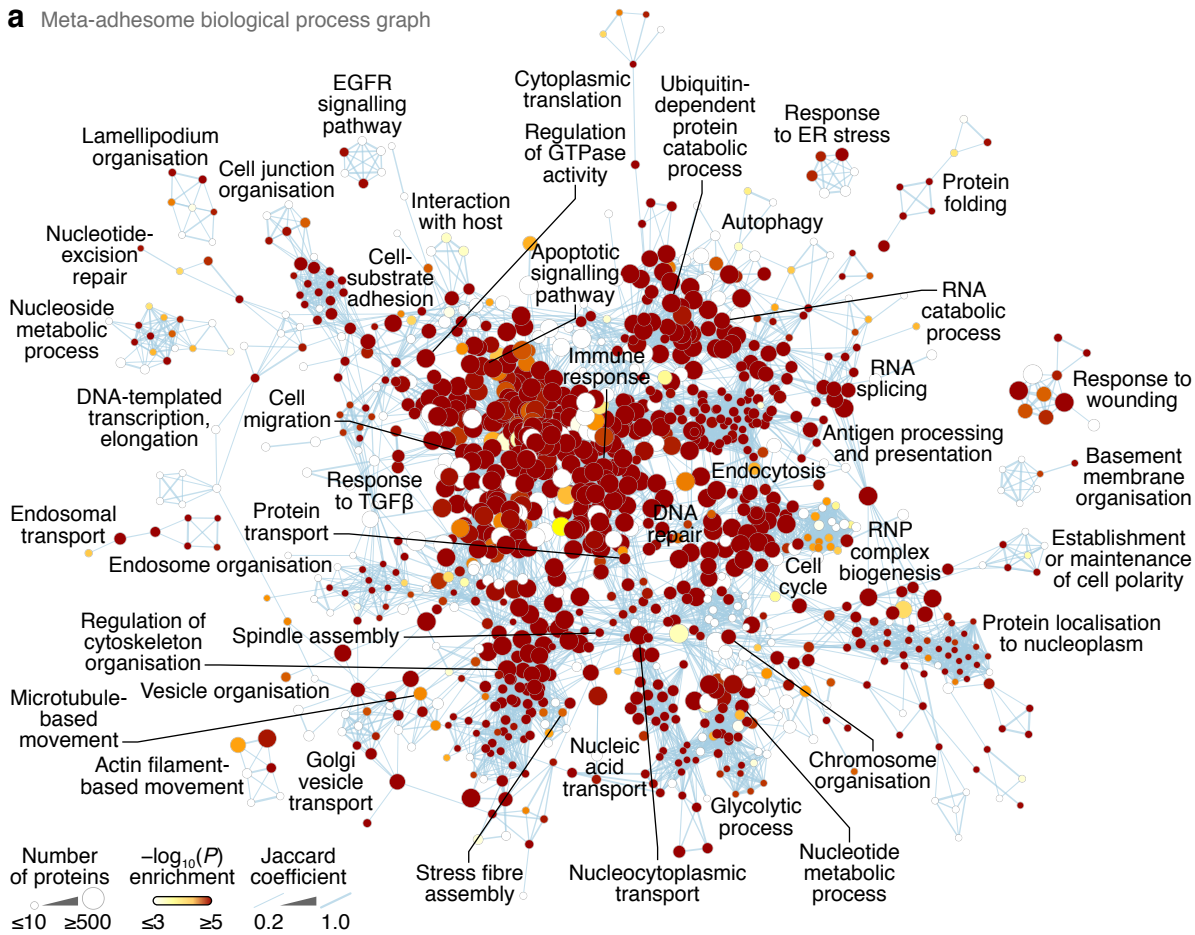
Supplementary Figure 3

Supplementary Figure 4

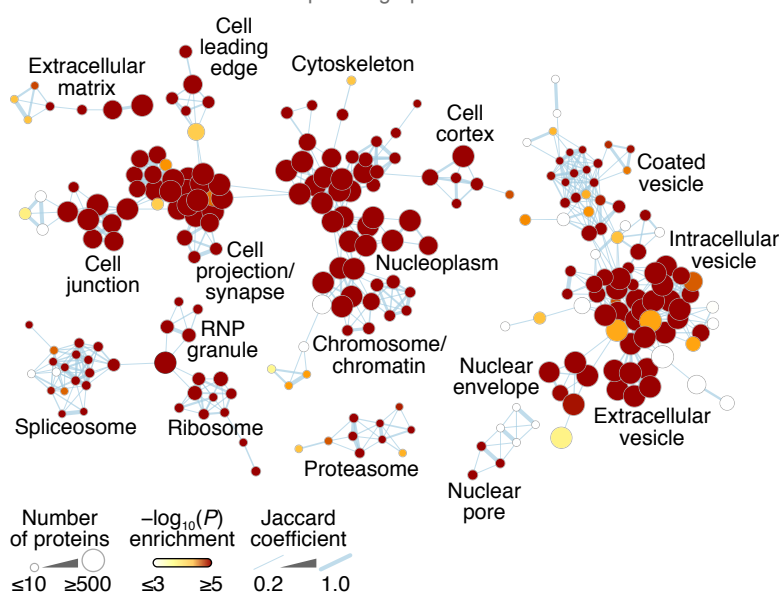
Supplementary Figure 5

Supplementary Table 1

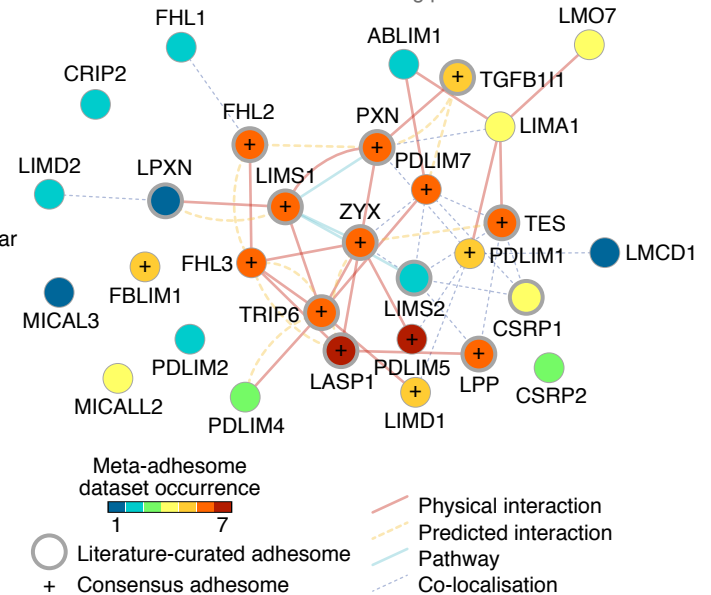
a Meta-adhesome biological process graph



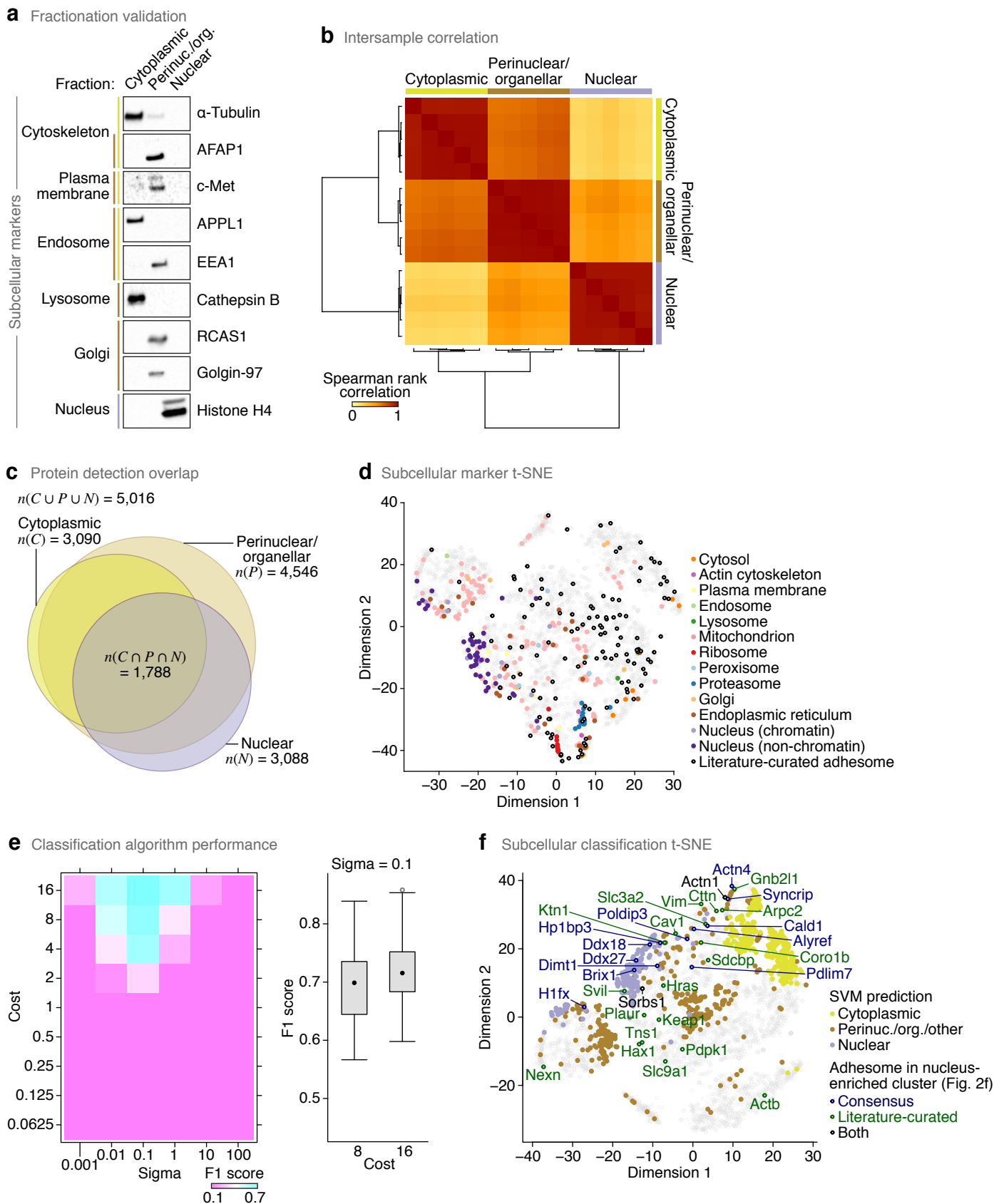
b Meta-adhesome cellular component graph



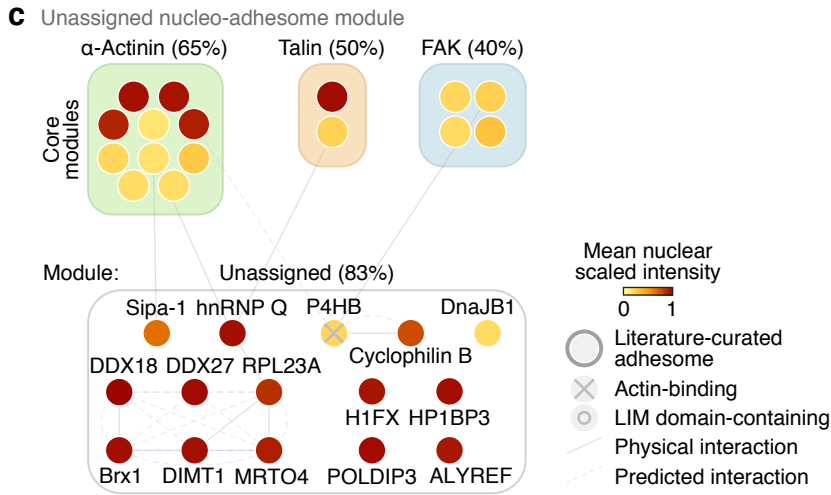
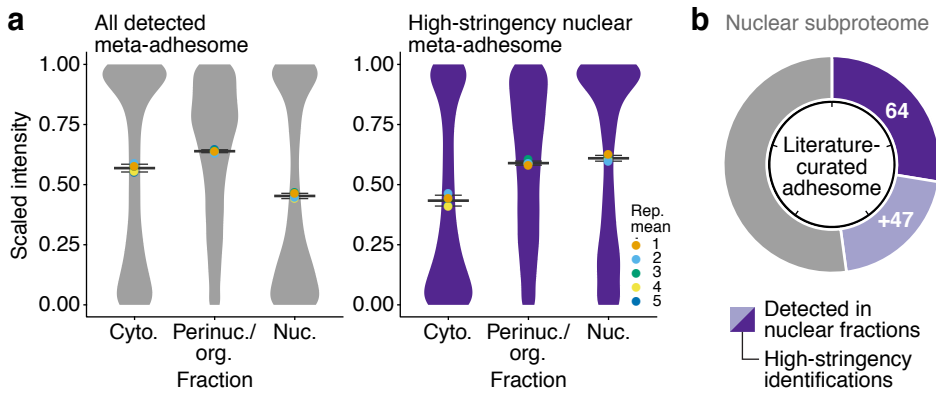
c Meta-adhesome LIM domain-containing protein network



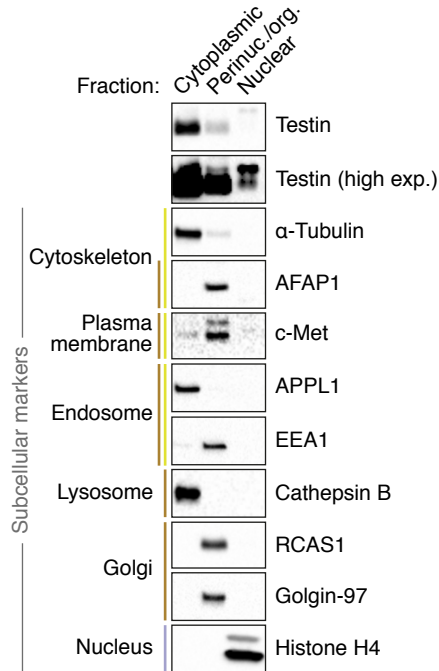
Supplementary Figure 1. Functional network analysis of the meta-adhesome. **(a, b)** Graph-based clustering of over-represented biological processes **(a)** and cellular components **(b)** in the meta-adhesome. Edge-weighted force-directed graphs were generated from gene-set memberships of enriched Gene Ontology terms ($P < 0.01$, one-sided hypergeometric test with Benjamini–Hochberg correction). Connected components with more than five nodes (enriched terms) are shown. Representative functional categories summarising selected clusters of enriched terms are labelled. Node (circle) size represents the number of meta-adhesome proteins annotated for each enriched term; node fill colour represents the significance of enrichment. Edge (line) weight is proportional to the overlap of gene-set membership of connected nodes (Jaccard coefficient ≥ 0.2). Selected clusters are indicated in Fig. 1c, d. **(c)** Functional association network of LIM domain-containing proteins in the meta-adhesome ($P = 6.04 \times 10^{-8}$, one-sided hypergeometric test with Benjamini–Hochberg correction). Node fill colour represents frequency of occurrence in meta-adhesome datasets; thick node borders indicate representation in the literature-curated adhesome. Edges represent reported associations. Proteins (nodes) are annotated with gene names for clarity.



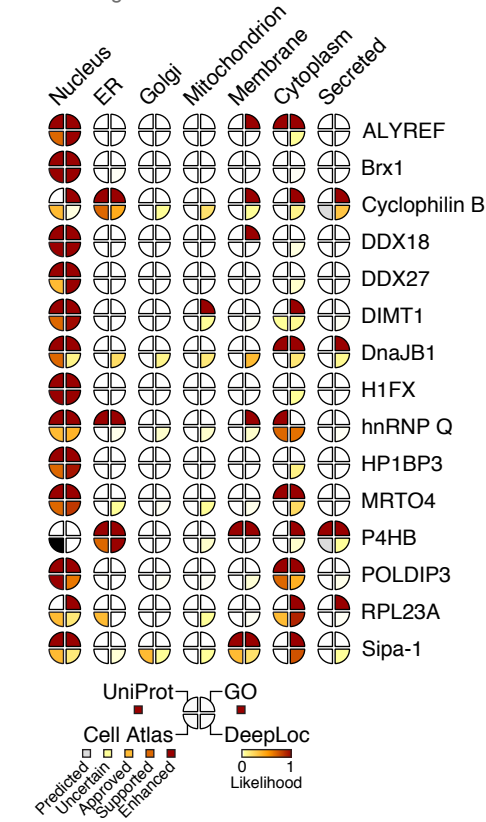
Supplementary Figure 2. Mass spectrometric analysis of the subcellular proteomes of SCC cells. **(a)** Subcellular fractionation of SCC cells. Markers for given subcellular locations are indicated. Immunoblots are representative of three independent experiments. See also Fig. 2b. **(b)** Hierarchical cluster analysis of Spearman rank correlation coefficients for all pairwise sample comparisons based on proteins quantified by LC-MS/MS in at least four out of five biological replicates. **(c)** Area-proportional Euler diagram of numbers of proteins detected and quantified in different subcellular fractions. Cardinalities of selected sets are defined, where C = cytoplasmic, P = perinuclear/organelle and N = nuclear fractions. **(d)** t-SNE map of the subcellular proteomes (Supplementary Data 2) annotated with curated subcellular markers and literature-curated adhesome proteins. See also Fig. 2g. **(e)** SVM algorithm performance. Heat map of the average F1 scores (harmonic means of precision and recall) for the full range of cost–sigma parameter pairs tested (left). Box plot summarising the distribution of F1 scores for the best parameter pairs (right). **(f)** t-SNE map of the subcellular proteomes annotated with subcellular SVM-based classification (Supplementary Data 2) and consensus and literature-curated adhesome proteins identified in the nucleus-enriched cluster by hierarchical clustering (Fig. 2f) (proteins are labelled with gene names for clarity). Source data are provided as a Source Data file.



d Testin fractionation validation

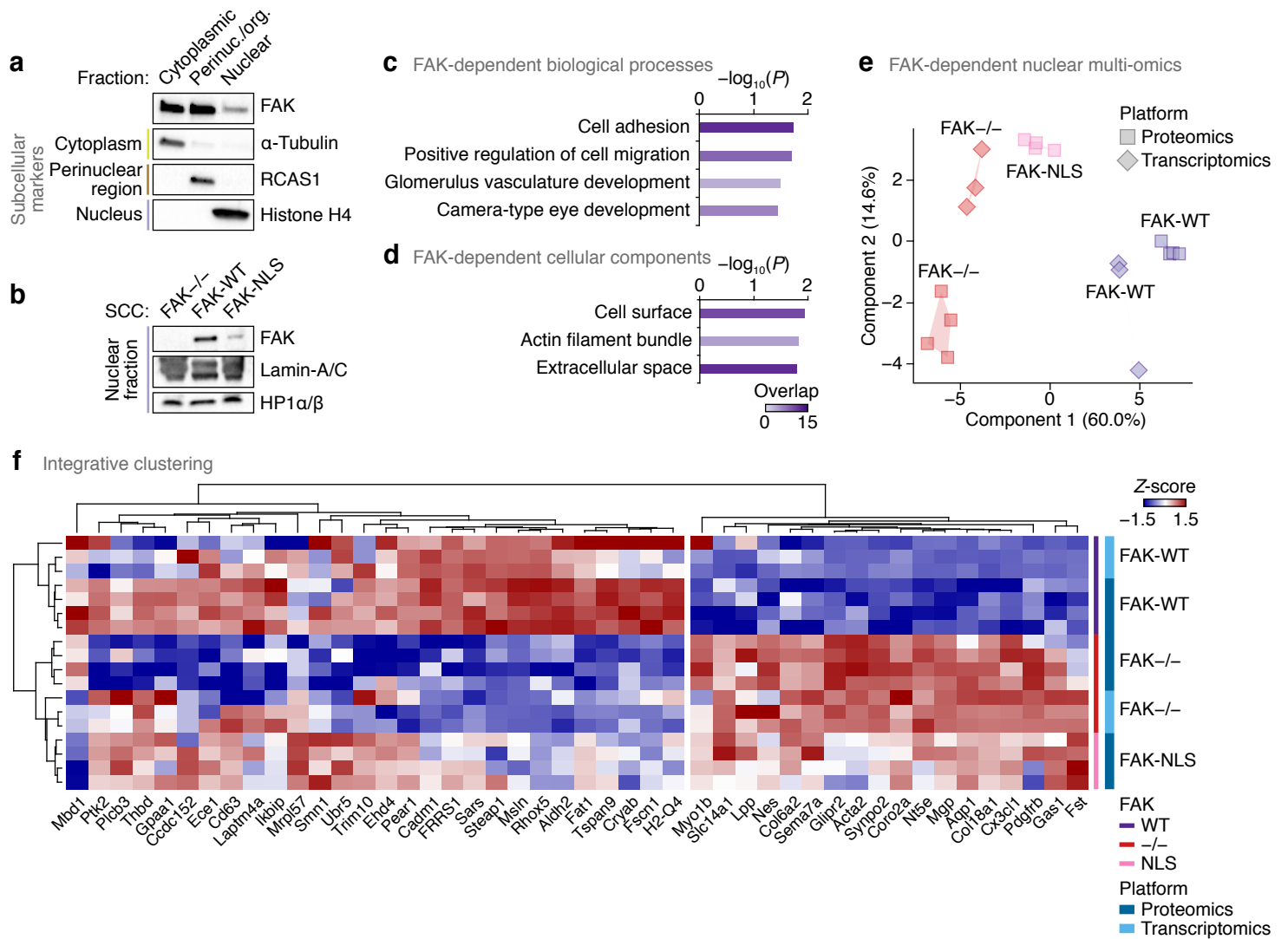


e Unassigned module subcellular annotation

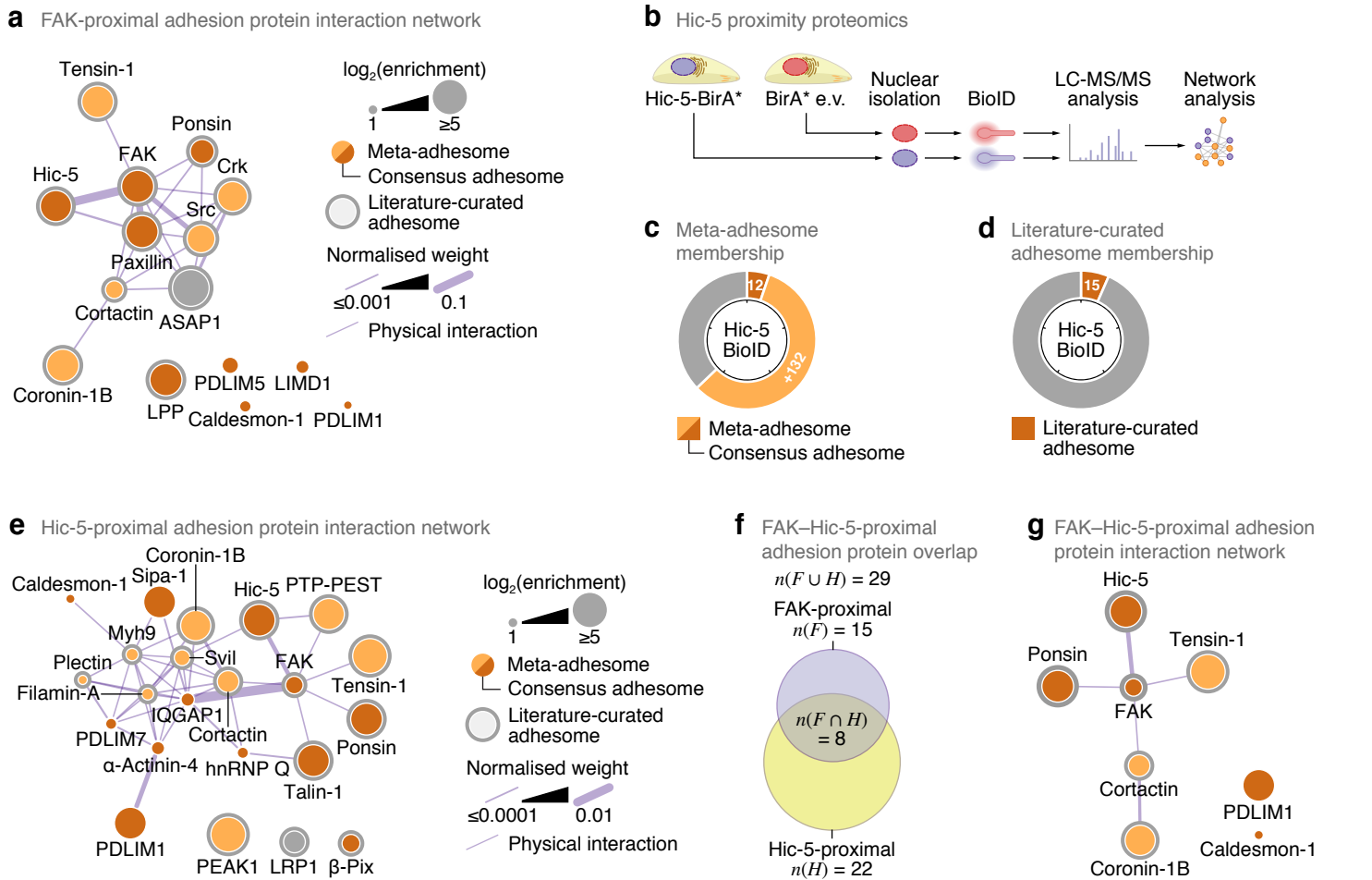


Supplementary Figure 3. See next page for caption.

Supplementary Figure 3. Network analysis of a nucleo-adhesome. (a) Relative quantification of meta-adhesome proteins across SCC subcellular fractions. Distributions of scaled intensities are shown for all detected meta-adhesome proteins (left) and meta-adhesome proteins identified in nuclear fractions with high stringency (right). Black bars, fraction mean (thick bar) \pm s.d. (thin bars); silhouette, probability density. Independent biological replicate (rep.) means are indicated by coloured circles ($n = 9,425$ and $6,060$ proteins for all detected meta-adhesome proteins and high-stringency nuclear meta-adhesome proteins, respectively, from $n = 5$ independent biological replicates). (b) Proportion of the literature-curated adhesome quantified in SCC cell nuclear fractions. High-stringency proteins (dark purple segments) were identified in all five biological replicate experiments and quantified with a $>5\%$ fraction of the cellular pool; light purple segments indicate additional proteins detected in nuclear fractions. Tick marks indicate 20% increments. (c) Network model of high-stringency nuclear proteins present in the consensus adhesome but not associated with an assigned module (termed unassigned). Coverage of the module is indicated in parentheses. Node (circle) fill colour represents the mean scaled intensity of nuclear fraction replicates; thick grey node borders indicate representation in the literature-curated adhesome. Edges (lines) represent reported interactions. (d) Subcellular fractionation of SCC cells. Markers for given subcellular locations are indicated. Immunoblots are representative of three independent experiments. (e) Subcellular annotation plot of high-stringency nuclear proteins in the unassigned nucleo-adhesome module. Segments are coloured according to annotation in the UniProt knowledgebase, annotation in the Gene Ontology (GO) cellular component domain, Cell Atlas reliability class and DeepLoc prediction likelihood as indicated. For Cell Atlas annotation, nucleus represents the highest-scoring reliability class of annotated nuclear subregions (e.g. nucleoli, nucleoplasm, nuclear membrane); black segment indicates annotation in mouse cells only (no corresponding reliability class). Source data are provided as a Source Data file.



Supplementary Figure 4. Mass spectrometric analysis of the FAK-dependent nuclear proteome. **(a)** Subcellular fractionation of SCC cells. Markers for given subcellular locations are indicated. Immunoblots are representative of three independent experiments. **(b)** Nuclear fractionation of SCC cells that express FAK-WT or FAK-NLS or do not express FAK (FAK^{-/-}). Lamin-A/C and heterochromatin protein 1 α/β (HP1 α/β) are nuclear markers. Immunoblots are representative of two independent experiments. **(c, d)** Over-representation analysis of Gene Ontology biological processes **(c)** and cellular components **(d)** in the FAK-dependent nuclear proteome. Representative terms were determined using affinity propagation. Purple shading intensity indicates size of subproteome overlap with respective gene sets ($P < 0.05$, one-sided hypergeometric test with Benjamini–Hochberg correction). **(e)** Principal component analysis of integrated FAK-dependent nuclear proteomic and FAK-dependent transcriptomic (multi-omic) data. **(f)** Integrative cluster analysis of the FAK-dependent nuclear multi-omic data. Proteins are labelled with gene names for clarity. Sample (row) dendrogram also shown in Fig. 4i. Source data are provided as a Source Data file.



Supplementary Figure 5. Mass spectrometric analyses of the FAK- and Hic-5-proximal nuclear interactomes. (a) Interaction network analysis of specific FAK-proximal nuclear proteins present in the consensus or literature-curated adhesomes. (b) Workflow for mass spectrometric characterization of Hic-5-proximal nuclear proteins in SCC cells. (c, d) Proportions of specific Hic-5-proximal proteins ($P < 0.05$, one-sided Student's t -test with FDR correction) quantified in the meta-adhesome, including the consensus adhesome (c), and the literature-curated adhesome (d). Tick marks indicate 20% increments. (e) Interaction network analysis of specific Hic-5-proximal nuclear proteins present in the consensus or literature-curated adhesomes. Myh9, myosin-9 (also known as myosin-IIa); Svitl, supervillin. (f) Area-proportional Euler diagram of numbers of consensus or literature-curated adhesome proteins identified in FAK- and Hic-5-proximal nuclear interactomes. Cardinalities of sets are defined, where F = FAK-proximal and H = Hic-5-proximal nuclear adhesion proteins. (g) Interaction network analysis of the intersection of specific FAK- and Hic-5-proximal nuclear proteins present in the consensus or literature-curated adhesomes. Key as for e. For a, e and g, networks were constructed from reported physical protein interactions (edges) and weighted by linear regression according to evidence of protein co-association in the context of that network. Proteins are clustered according to connectivity; unconnected proteins also shown.

Supplementary Table 1 RNA-Seq data statistics.

SCC	Biological replicate	RIN	Clean reads ($\times 10^6$)	Clean bases ($\times 10^8$)	Phred quality score (Q) 20 (%)	GC content (%)
FAK-WT	1	9.6	70.3	70.3	98.2	50.9
FAK-WT	2	9.6	84.7	84.7	98.1	51.0
FAK-WT	3	9.7	78.1	78.1	98.2	50.8
FAK-/-	1	9.7	75.9	75.9	98.1	51.1
FAK-/-	2	9.6	57.6	57.6	98.1	51.0
FAK-/-	3	9.7	72.9	72.9	98.0	53.1

RIN RNA integrity number.