Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

1. Sample size
   Describe how sample size was determined.
   No statistical methods were used to determine sample size

2. Data exclusions
   Describe any data exclusions.
   No data was excluded from the analysis.

3. Replication
   Describe whether the experimental findings were reliably reproduced.
   Figure 1c, d: Data are representative of two independent mESC glycoproteomics experiments (technical and biological replicates) with similar results.
   Figure 2: in all panels, data are representative of two independent mESC glycoproteomics experiments (technical and biological replicates) with similar results and one multiplexed hESC glycoproteomics experiment, analysing two technical and two biological replicates, with similar results for the 4 individual hESC samples.
   Figure 3c, d: Data are representative of two independent mESC glycoproteomics experiments (technical and biological replicates) with similar results.
   Figure 4a, b, c, d: Data are shown as mean ± SD of triplicate cultures. Experiments were repeated a minimum of three times with similar results.
   Figure 4e: Each image is representative of six images of three independent experiments.

4. Randomization
   Describe how samples/organisms/participants were allocated into experimental groups.
   Not applicable.

5. Blinding
   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.
   Investigators were not blinded.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.
6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
- A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- A statement indicating how many times each experiment was replicated
- The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
- A description of any assumptions or corrections, such as an adjustment for multiple comparisons
- The test results (e.g. P values) given as exact values whenever possible and with confidence intervals noted
- A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
- Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

7. Software

Policy information about availability of computer code

Describe the software used to analyze the data in this study.

All MS/MS data were processed and analysed using Xcalibur Version 2.2.48 (Thermo) and Proteome Discoverer 1.4 (PD 1.4.0.288, Thermo). All Proteome Discoverer 1.4 Nodes and in silico workflows are available for download at http://ms.imp.ac.at/?goto=kassonade.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). Nature Methods guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

Not restrictions on availability apply.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

Primary antibodies specific for IGF2R (anti M6PR (cation independent) antibody (MEM-238), abcam ab8093) were validated by the supplier and confirmed by specific reactivity against IGF2R in an over-expressing versus a IGF2R deficient cell line (i.e. SCC-VII/IGF2R, SCC-VII/IGF2R* versus parental SCC-VII).

The anti Oct-3/4 antibody (Clone 40/Oct-3 (RUO), BD Transduction Laboratories, No.611203, 1:300) and anti-SSEA-1 (CD15, Lewis X; PE conjugated, clone: MC-480, ebioscience 12-8813-41, 1:300) antibody were validated by the suppliers.
10. Eukaryotic cell lines  

<table>
<thead>
<tr>
<th>a. State the source of each eukaryotic cell line used.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Murine AN3-12 ESC lines have been generated in our lab. Mouse SCC-VII cells were originally isolated by Dr. R. P. Hill (Ontario Cancer Institute, Toronto, Canada) and obtained and provided to this study by Prof. Lukas Mach (BOKU, Vienna). HEK293FT were obtained from an in-house source. H9 human ESCs were obtained directly from the supplier (WiCell, Wisconsin, U.S.A).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>b. Describe the method of cell line authentication used.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Murine AN3-12 ESC lines have been generated in our lab and are fully characterized and authenticated as described (Elling U et al. (2011) Cell Stem Cell 9, 563-574). Mouse SCC-VII cells were previously described (Probst O. et al. (2013) Biochem. J. 451, 91-99) and authenticated by the absence of endogenous Igf2r, and presence of a functional transgene (i.e. IGF2R, IGF2R*). HEK293FT cells were functionally authenticated by positive selection for pCMVSPORT6Tag.neo expression. H9 human ESCs were received directly from the supplier (WiCell) and thus not further authenticated.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>c. Report whether the cell lines were tested for mycoplasma contamination.</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cell lines were tested negative for Mycoplasm contamination.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.</th>
</tr>
</thead>
<tbody>
<tr>
<td>No cell line listed by ICLAC was used.</td>
</tr>
</tbody>
</table>

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

Policy information about studies involving human research participants

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

[Not applicable]

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

[Not applicable]
Flow Cytometry Reporting Summary

Form fields will expand as needed. Please do not leave fields blank.

Data presentation

For all flow cytometry data, confirm that:

☑ 1. The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ 2. The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a ‘group’ is an analysis of identical markers).

☐ 3. All plots are contour plots with outliers or pseudocolor plots.

☒ 4. A numerical value for number of cells or percentage (with statistics) is provided.

Methodological details

5. Describe the sample preparation. Please see "Supplementary Information", section "Immunofluorescence"

6. Identify the instrument used for data collection. FACS LSR Fortessa is a 4 laser flow cytometer with 16 fluorescence detectors controlled by FACSDiva software

7. Describe the software used to collect and analyze the flow cytometry data. FACSDiva and FlowJo software were used to collect and analyze flow cytometry data. Further data analysis was performed using the statistical software package Prism (version 7)

8. Describe the abundance of the relevant cell populations within post-sort fractions. Flow cytometry was used for quantification purposes only (i.e. no post-sorting fractions were collected)

9. Describe the gating strategy used. For all experiments FSC-A/ SSC-A gates of the starting cell population were used to discriminate between viable cells and cell debris. Singlet and doublet cells were discriminated using FSC-A/ FSC-W gating. Isotype control stained cells were used to distinguish between background staining and specific antibody staining. For transgenic cell lines expressing fluorescent proteins, parental cell lines not expressing the respective fluorophores were used as negative controls.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information. ☐