Suppl. Figure 1. Assay validation and multi-parametric profiling of endocytic phenotypes with QMPIA

(a-c) Validation of the assay and QMPIA. High resolution images of HeLa cells transfected with siRNA silencing CLTC (a), EGFR (b) and TFRC (c) prior to internalization of fluorescently labelled EGF and Tf for 10’. Pseudo-colouring is as in Fig. 1a.

(d-e) Multi-parametric profiles of the 46 most specific parameters (listed in Suppl. Table I and Suppl. Material) are calculated and plotted as shown in Fig. 1d corresponding to the cells in (a-c).
Suppl. Figure 2. Assay validation and multi-parametric profiling of endocytic phenotypes with QMPIA

(a-f) Validation of the assay and QMPIA. High resolution images of HeLa cells Mock treated or transfected with siRNA silencing Dynamin-2 (DNM2) (b), EEA1 (EEA1) (c), Rabenosyn-5 (ZFYVE20) (d), VPS45 (VPS45) (e) and Rabkyrin-5 (ANKFY1) (f). Cells internalized fluorescently

www.nature.com/nature
labelled EGF and Tf for 10’ and were subsequently fixed. Pseudo-colouring is as in Fig. 1a.

(g–l) Corresponding multi-parametric profiles are shown. The 46 most specific parameters (listed in Suppl. Table I and Suppl. Material) are aligned on the X axis and normalized z-values are plotted on the Y axis as shown in Fig. 1d. Only parameter groups (e.g. G1, G2 etc.) are indicated. The horizontal bars on the top indicate the markers to which the parameters refer to.

Suppl. Figure 3. Correlation of individual confocal planes to the maximal projection of the entire cell

Ten confocal planes with 0.5 μm steps were acquired for a small set of 27 siRNAs. The average Pearson correlation coefficient between each individual plane and the maximal projection was calculated across all parameters of the QMPIA and plotted in the bar graph. The panel shows that the correlation between the individual planes and their maximal projection is in the range 0.47-0.79 and that plane 2 (Pl_2; the second plane from the bottom, which has been used for all further imaging in the screen) displays the maximal correlation (0.79±0.029, higher than the correlation between experiments 0.58±0.04). The correlations between plane 2 and the other neighbouring planes plane 1±plane 6 are in the range of 0.6-0.85 (not shown).
Suppl. Figure 4

Suppl. Figure 4. Example of a randomly chosen hit gene multi-parametric profile and its corresponding si/esiRNAs profiles.

A total of 8 si/esiRNAs were analysed targeting the gene. (a-h) The profiles of the individual si/esiRNAs targeting a randomly chosen hit gene are compared to the corresponding gene profile.
profile (red line). The si/esiRNA profiles have been colour coded depending on the consistency
(red consistent, blue inconsistent,) on each individual parameter to the gene profile. On each
si/esiRNA profile, red arrows indicate the consistent parameters and blue arrows the
inconsistent parameters. Panel (i) shows the gene profile with error bars for each parameter
overlaid to all si/esiRNA profiles. Encircled are two regions of the profile showing high (red oval)
and low (blue oval) consistency between si/esiRNAs.

Three points are readily visible:
1) Each individual si/esiRNA has both regions of consistency and regions of inconsistencies in
the profile. This illustrates the impossibility of discretely selecting the number of “matching”
si/esiRNAs using thresholds.

2) In the consistency region (red oval, EGF mean concentration and content) the phenotype is
“supported” by 6 out of 8 si/esiRNAs whereas in the inconsistency region (blue oval) there are
contrasting si/esiRNAs responses on Tfn parameters where the positive effects are neutralized
by negative effects. In the latter case, a mono-parametric assay with 2-3 siRNAs has high
probability to hit a false positive by selecting 2 or 3 matching siRNAs.

3) The reproducibility of the gene profile as “average” profile of all si/esiRNA profiles is visible
from the comparison between “signal” amplitude and the error bars. In the region of
consistency the amplitude of the signal is > 4 SEM.

It can be noticed how the availability of 1) quantitative measurements and 2) multiple
parameters allows the averaging out of the discrepancies and the selection of specific
phenotypic effects supported by the majority of si/esiRNAs screened.
**Suppl. Figure 5. Screening plates design**

Design of the screening plates is illustrated. Plate design for the Ambion library (a), for the Qiagen library (b), for the esiRNA library (c). In all cases the wells containing the si/esiRNAs from the libraries are indicated in yellow, the Mock (no si/esiRNAs) wells are indicated in green and the untreated samples (without both transfection reagent and si/esiRNAs) are indicated in grey. Each plate contains transfection efficiency controls (si/esiRNAs targeting INCENP and TPX2, indicated in red), QMPIA controls (si/esiRNAs yielding well defined multi-parametric profiles: VPS45A, RIN2, ZFYVE20, EEA1, FLT-4, DTYMK,  for the Ambion and Qiagen libraries, and VPS45A, INPP4A(esi), ZFYVE20, EEA1(esi), FLT-4(esi), FLT-4, STK35(esi), STK35 for the esiRNA library indicated in violet and light blue) and negative control si/esiRNAs (Ambion5 and Ambion6 for the Ambion library, Luc-GL2, Luc-GL3 and Allstar for the Qiagen library and R-Luc1(esi), F-Luc2(esi) and Allstar for the esiRNA library).
Suppl. Figure 6. Score selection criteria

The criteria for the selection of the positive genes are schematically illustrated. The probability $\chi^2$ was calculated as the maximal value between $\chi^2_{total\_profile}$ ($1.0 - p_{value}$ of the gene profile $\chi^2$), $\chi^2_{EGF}$ ($1.0 - p_{value}$ of the gene profile $\chi^2$ considering only the parameters for EGF) and $\chi^2_{Tfn}$ ($1.0 - p_{value}$ of the gene profile $\chi^2$ considering only the parameters for Tfn). All genes with a probability $\chi^2 \geq 0.95$ (“strong” phenotypes) were directly included in the list of positive whereas genes with a probability $\chi^2 < 0.95$ were divided in two groups. The genes with a probability $\chi^2 < 0.75$ were considered negative and for the genes with a probability $\chi^2 \geq 0.75$ the pheno-score probability was calculated (see Suppl. Material). All genes with a pheno-score probability ($1.0 - p_{value}$) $\geq 0.95$ (“mild” but specific phenotypes) were included in the list of positive and the genes with pheno-score probability ($1.0 - p_{value}$) $< 0.95$ were considered negative.
Suppl. Fig. 7. Protein silencing validation for a small set of siRNAs

(a-c) Levels of protein silencing in HeLa cells transfected with the indicated siRNAs targeting the indicated genes 72h after transfection measured by Western blotting. Six different siRNAs (IDs are indicated) were used to silence each of the 3 indicated genes of the human genome (EEA1, ZFYVE20 and ANKFY1). Relative band intensities are plotted in the bar graphs. Band intensities were normalized in percent of control (Mock transfected cells). Stars indicate lethal siRNAs.
Suppl. Figure 8. Examples of phenotypic clusters

Phenotypic clusters identified by mean shift clustering (see Suppl. Material) of the genomic set of gene profiles are described in Fig. 3k. The profiles of all fourteen main phenotypic cluster groups are shown in Fig. 3. Multi-parametric profiles representing the mode profile of each cluster-group (cluster-group profiles) were plotted as described in Fig. 1. This panel shows examples of anti-correlated cluster-groups: cluster-group 1 vs. 2 (a) and cluster-group 3 vs. 4 (b). The amplitude of the phenotype vectors has been normalized to 1. Light microscopy images (nuclei and cytoplasm in blue, EGF in red and Tfn in green) representative of cluster groups with opposite phenotypes are shown for cluster 1 (c) vs. cluster 2 (d) and cluster 3 (e) vs. cluster 4 (f).
Suppl. Figure 9

Suppl. Figure 9. Surface receptors staining

All si/esiRNAs targeting the genes of the Notch and TGFβ signalling pathways of cluster-group 8 were screened along with a set of 1,034 randomly selected si/esiRNAs to detect changes in the level of surface receptors. Fluorescently labelled Alexa-488-EGF and Alexa-647-Tfn were added to cells for 30 min on ice before fixation. Surface receptor signals amplification was performed by staining with antibodies anti-Alexa-488 and TFRC (ectodomain) without membrane permeabilization. (a) Fluorescent microscopy images of Mock-treated HeLa cells. Fluorescent
signals from Alexa-488-EGF and anti-Alexa-488 antibodies are pseudocolored in red, Alexa-647-Tfn and anti-Transferrin receptor antibodies in green and DAPI-SYTO 42 in blue.

(b-c) Average (mode basis; see Suppl. Material) normalized z-values of the total fluorescence intensity normalized by the area covered by cells for the indicated genes and assay controls are plotted in the bar graphs. “Control” indicates Mock treated cells; “Control-No primary Ab”, mock-treated cells incubated in the absence of primary antibody; “No-EGF/Tfn”, mock-treated cells where no fluorescently labelled ligands were added; “No-EGF/Tfn-No primary Ab”, mock-treated cells where no fluorescently labelled ligands and no primary antibody were added; “cold EGF/Tfn”, mock-treated cells where incubation with the fluorescently labelled ligands was performed in the presence of 100 fold excess of unlabelled EGF and Tfn. No significant alteration of the steady-state amounts of EGFR and TFRC on the cell surface was detectable for the assayed genes.
Suppl. Figure 10. Several genes required for Tfn uptake are not required for EGF uptake

The unexpected low degree of correlation between EGF and Tfn uptake (0.15; Fig. 4c) suggests considerable diversity between the components required for the uptake of the two types of cargo. Microscopic images of cells knocked-down for the β-subunit of the AP-2 complex are presented in (b) in comparison to control cells (Mock; a) as example. Nuclei and cellular cytoplasm (blue), EGF (red) and Tfn (green).