

This is a repository copy of *A quantitative mass spectrometry-based approach to monitor the dynamics of endogenous chromatin-associated protein complexes*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/151263/>

Version: Published Version

---

**Article:**

Papachristou, EK, Kishore, K, Holding, AN [orcid.org/0000-0002-8459-7048](https://orcid.org/0000-0002-8459-7048) et al. (12 more authors) (2018) A quantitative mass spectrometry-based approach to monitor the dynamics of endogenous chromatin-associated protein complexes. *Nature Communications*. 2311. ISSN 2041-1723

<https://doi.org/10.1038/s41467-018-04619-5>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

ARTICLE

DOI: 10.1038/s41467-018-04619-5

OPEN

# A quantitative mass spectrometry-based approach to monitor the dynamics of endogenous chromatin-associated protein complexes

Evangelia K. Papachristou<sup>1</sup>, Kamal Kishore<sup>1</sup>, Andrew N. Holding<sup>1</sup>, Kate Harvey<sup>2</sup>, Theodoros I. Roumeliotis<sup>3</sup>, Chandra Sekhar Reddy Chilamakuri<sup>1</sup>, Soleilmane Omarjee<sup>1</sup>, Kee Ming Chia<sup>2</sup>, Alex Swarbrick<sup>2,4</sup>, Elgene Lim<sup>2,4</sup>, Florian Markowitz<sup>1</sup>, Matthew Eldridge<sup>1</sup>, Rasmus Siersbaek<sup>1</sup>, Clive S. D'Santos<sup>1</sup> & Jason S. Carroll<sup>1</sup>

Understanding the dynamics of endogenous protein-protein interactions in complex networks is pivotal in deciphering disease mechanisms. To enable the in-depth analysis of protein interactions in chromatin-associated protein complexes, we have previously developed a method termed RIME (Rapid Immunoprecipitation Mass spectrometry of Endogenous proteins). Here, we present a quantitative multiplexed method (qPLEX-RIME), which integrates RIME with isobaric labelling and tribrid mass spectrometry for the study of protein interactome dynamics in a quantitative fashion with increased sensitivity. Using the qPLEX-RIME method, we delineate the temporal changes of the Estrogen Receptor alpha (ER $\alpha$ ) interactome in breast cancer cells treated with 4-hydroxytamoxifen. Furthermore, we identify endogenous ER $\alpha$ -associated proteins in human Patient-Derived Xenograft tumours and in primary human breast cancer clinical tissue. Our results demonstrate that the combination of RIME with isobaric labelling offers a powerful tool for the in-depth and quantitative characterisation of protein interactome dynamics, which is applicable to clinical samples.

<sup>1</sup>Cancer Research UK Cambridge Institute, University of Cambridge, Robinson Way, Cambridge CB2 0RE, UK. <sup>2</sup>Garvan Institute of Medical Research, Darlinghurst, Sydney NSW 2010, Australia. <sup>3</sup>Wellcome Trust Sanger Institute, Wellcome Genome Campus, Cambridge CB10 1SA, UK. <sup>4</sup>St Vincent's Clinical School, UNSW, Sydney NSW 2052, Australia. These authors contributed equally: Kamal Kishore, Andrew N. Holding. Correspondence and requests for materials should be addressed to R.S. (email: [Rasmus.Siersbaek@cruk.cam.ac.uk](mailto:Rasmus.Siersbaek@cruk.cam.ac.uk)) or to C.S.D'S. (email: [Clive.D'Santos@cruk.cam.ac.uk](mailto:Clive.D'Santos@cruk.cam.ac.uk)) or to J.S.C. (email: [Jason.Carroll@cruk.cam.ac.uk](mailto:Jason.Carroll@cruk.cam.ac.uk))

Deciphering the role and the organisation of dynamically regulated protein networks is critical for the accurate molecular characterisation of biological systems<sup>1</sup>. Over the last decade, the advancements made in mass spectrometry-based proteomics have enabled the rapid analysis of complex protein samples obtained from co-immunoprecipitation assays, providing a powerful tool for the study of protein interactions and protein complexes<sup>2</sup>. In this regard, the first systematic efforts to generate human protein interactome maps using yeast two-hybrid<sup>3–5</sup> have been recently complemented by studies utilising large-scale Affinity Purification followed by Mass Spectrometry analysis (AP-MS)<sup>6,7</sup>. Additionally, the integration of AP-MS with quantitative approaches has enabled the study of stoichiometric changes in protein complexes<sup>8</sup>. More recently, the use of chemical crosslinking combined with mass spectrometry has provided information about endogenous protein assemblies in a proteome-wide scale<sup>9</sup>.

Gene regulation relies on the coordinated action of transcription factors and co-regulator complexes that control transcriptional activation at promoters or enhancers. To gain insight into the complex interactions between such regulators, the combination of Chromatin Immunoprecipitation (ChIP) with mass spectrometry has been used to study the composition of chromatin-associated complexes<sup>10–12</sup>. In line with this strategy we have previously developed RIME (Rapid Immunoprecipitation Mass spectrometry of Endogenous proteins)<sup>13</sup>, a method which has several advantages for the analysis of protein interactomes<sup>14</sup>. RIME provides a sensitive and rapid approach for the identification of protein complexes from low amounts of starting material and importantly involves purification of endogenous protein, rather than the use of exogenous tagged approaches.

In the present study, we have established a modified RIME assay to monitor the dynamics of chromatin-associated complexes using a quantitative multiplexed workflow (quantitative Multiplexed Rapid Immunoprecipitation Mass spectrometry of Endogenous proteins or qPLEX-RIME). Specifically, we combine RIME with isobaric labelling using Tandem Mass Tags (TMT-10plex)<sup>15,16</sup>, peptide fractionation and MultiNotch MS3 analysis<sup>17</sup>. This combination allows the simultaneous analysis of multiple conditions and biological replicates with high sensitivity in a single experiment. Additionally, we have developed a data analysis workflow termed quantitative Multiplexed analyzer (qPLEXanalyzer) that permits statistical analysis of the quantitative interactome data and the identification of differential interactions.

As a proof-of-concept, we apply the qPLEX-RIME method to discover the temporal changes of Estrogen Receptor alpha (ER $\alpha$ ) interactors in breast cancer cells treated with 4-hydroxytamoxifen (OHT) and to identify the ER $\alpha$  interactome in human patient-derived xenograft (PDX) tumours and in human breast cancer tissues. Our data demonstrate that the qPLEX-RIME method combines multiplexity with quantitative accuracy and increased sensitivity, to enable the in-depth characterisation of dynamic changes in chromatin-associated protein complexes *in vitro* and *in vivo*.

## Results

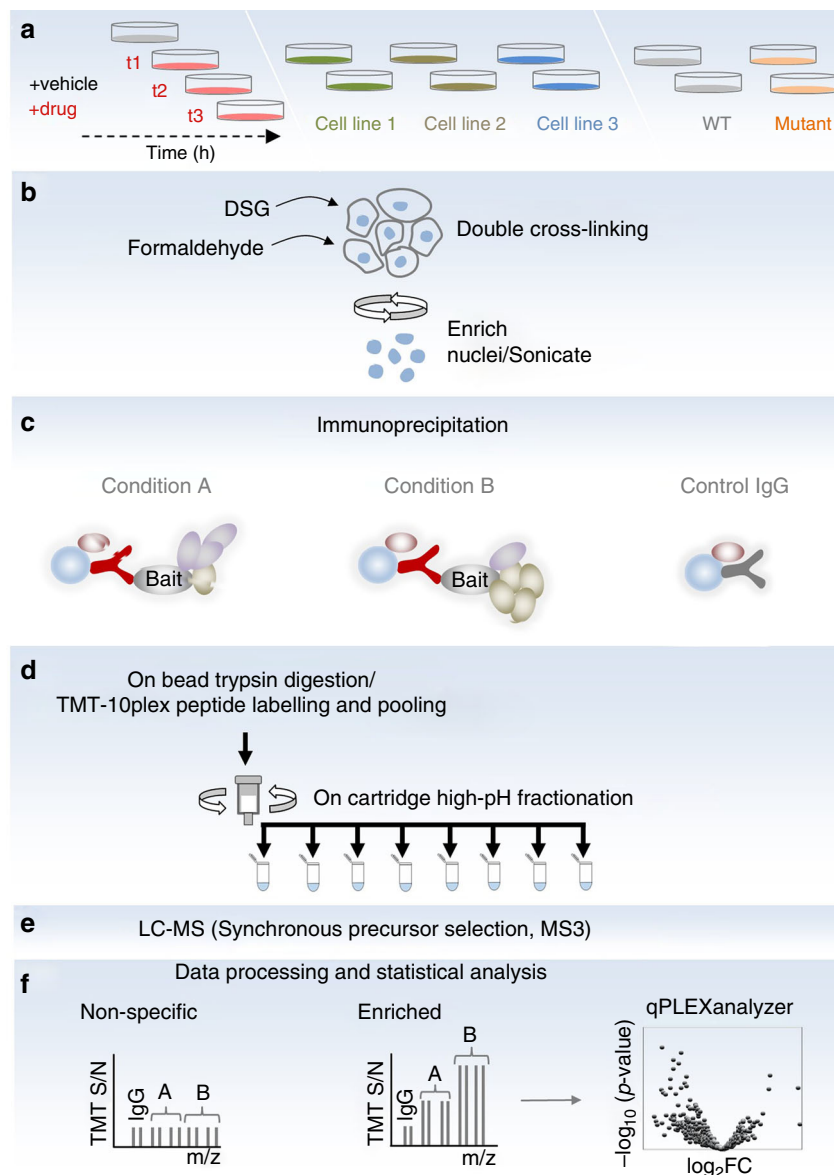
**The qPLEX-RIME workflow.** The qPLEX-RIME approach combines the RIME method<sup>13,14</sup> with multiplex TMT chemical isobaric labelling<sup>15,16</sup> to study the dynamics of chromatin-associated protein complexes. The workflow starts with a two-step fixation procedure using disuccinimidyl glutarate (DSG) and formaldehyde (FA) that has been previously applied in combination with ChIP assays to capture transient interactions more efficiently<sup>12,18</sup>. A specific antibody against the target protein is

used for immunoprecipitation, followed by proteolysis, TMT-10plex peptide labelling and fractionation. The main steps of the qPLEX-RIME method are shown in Fig. 1. The main utility of the qPLEX-RIME method is the quantification of changes in the composition of protein complexes in response to cell perturbation and/or in variable genomic backgrounds (e.g. different cell lines or mutated conditions) using multiple biological replicates in a single experiment. Also proteins that are significantly and specifically associated with the bait protein can be discovered in the same analysis using appropriate negative controls, such as IgG pull-downs. For the downstream data analysis, we have developed a comprehensive bioinformatics workflow (qPLEXanalyzer) that includes data processing, visualisation, normalisation and differential statistics. In addition to the qPLEXanalyzer R package, the complete qPLEX-RIME and full proteome data sets of this work are included in the qPLEXdata R package. Both packages can be found at (<https://doi.org/10.5281/zenodo.1237825>) and a detailed description of the pipeline and the applications is provided in Supplementary Notes 1 and 2.

**Characterisation of the ER $\alpha$  interactome in MCF7 cells.** We first applied qPLEX-RIME to assess whether we could successfully identify the ER $\alpha$  interactome in asynchronous MCF7 breast cancer cells. To this end, we performed ER $\alpha$  qPLEX-RIME pull-downs in five independent biological replicates. An equal number of matched IgG control samples were prepared. In this experiment we used single crosslinking with FA, to permit a comparison with previously published approaches<sup>13</sup>. In addition to the qPLEX-RIME, we included a standard non-quantitative ER $\alpha$  RIME experiment with matched IgG controls (Supplementary Data 1).

The qPLEX-RIME raw data processing quantified 2955 proteins across the multiplexed set of all positive and negative samples at peptide false discovery rate (FDR) <1% (Supplementary Data 2). To test the efficiency of the method in capturing and quantifying previously described ER $\alpha$ -associated proteins, we compiled a list of known ER $\alpha$  interactors from BioGRID<sup>19</sup> and STRING<sup>20</sup> resources. For BioGRID, we used only a subset of 386 proteins identified by high-throughput assays that are similar to the approach used here and for STRING we used only experimental associations (383 proteins, score >200). Noteworthy, only 37 proteins were common between the two reference subsets. The qPLEX-RIME method identified 295 (76%) and 171 (45%) of the known ER $\alpha$ -associated proteins from BioGRID and STRING, respectively, of which 225 (58%) and 154 (40%) showed positive enrichment at adj. *p*-value <0.1 (Limma moderated *t*-test) (Fig. 2a). Specifically, we found known co-regulators (e.g. EP300, NCOA3, CBP, NR1P1, TRIM24, GREB1, RAR $\alpha$ , NCOA2 and HDACs<sup>13,21–25</sup>), ER $\alpha$ -associated pioneer factors (e.g. FOXA1<sup>26</sup> and AP-2 $\gamma$ <sup>27</sup>), and putative pioneer factors (e.g. GATA-3<sup>28</sup>) with significant enrichment in the ER $\alpha$  samples (Fig. 2b).

ER $\alpha$  was one of the most significantly enriched proteins identified with 19 unique peptides (Fig. 2c), which is consistent with previously published ER $\alpha$  RIME experiments<sup>14</sup>. A comparison between the non-quantitative RIME and the qPLEX-RIME data showed that 302 of the 323 (93%) proteins identified as ER $\alpha$ -specific in the non-quantitative ER $\alpha$ -RIME pull-down analysis were also identified by qPLEX-RIME with significant enrichment over the IgG controls (mean fold-change of 2.5). Notably, the application of qPLEX-RIME achieved overall better peptide coverage for the overlapping ER $\alpha$ -associated proteins compared to the non-quantitative RIME method (Fig. 2d). Additionally, qPLEX-RIME identified 124 more known BioGRID and STRING interactors compared to the non-quantitative RIME



**Fig. 1** The qPLEX-RIME workflow. Proteins in treated and un-treated cell cultures at different time points or in variable genomic backgrounds (e.g. different cell lines or mutated conditions) are double-crosslinked and cell nuclei are isolated and sonicated (**a**, **b**). Target protein complexes are immunoprecipitated and subjected to on-bead trypsin digestion (**c**, **d**). The generated peptides are labelled using different TMT reagents and pooled in a single mixture, which is fractionated using Reversed-Phase cartridges (**d**). Peptide fractions are analysed with the MultiNotch MS3 method (**e**) followed by data processing and statistical analysis using novel analytic suite (qPLEXanalyzer) (**f**)

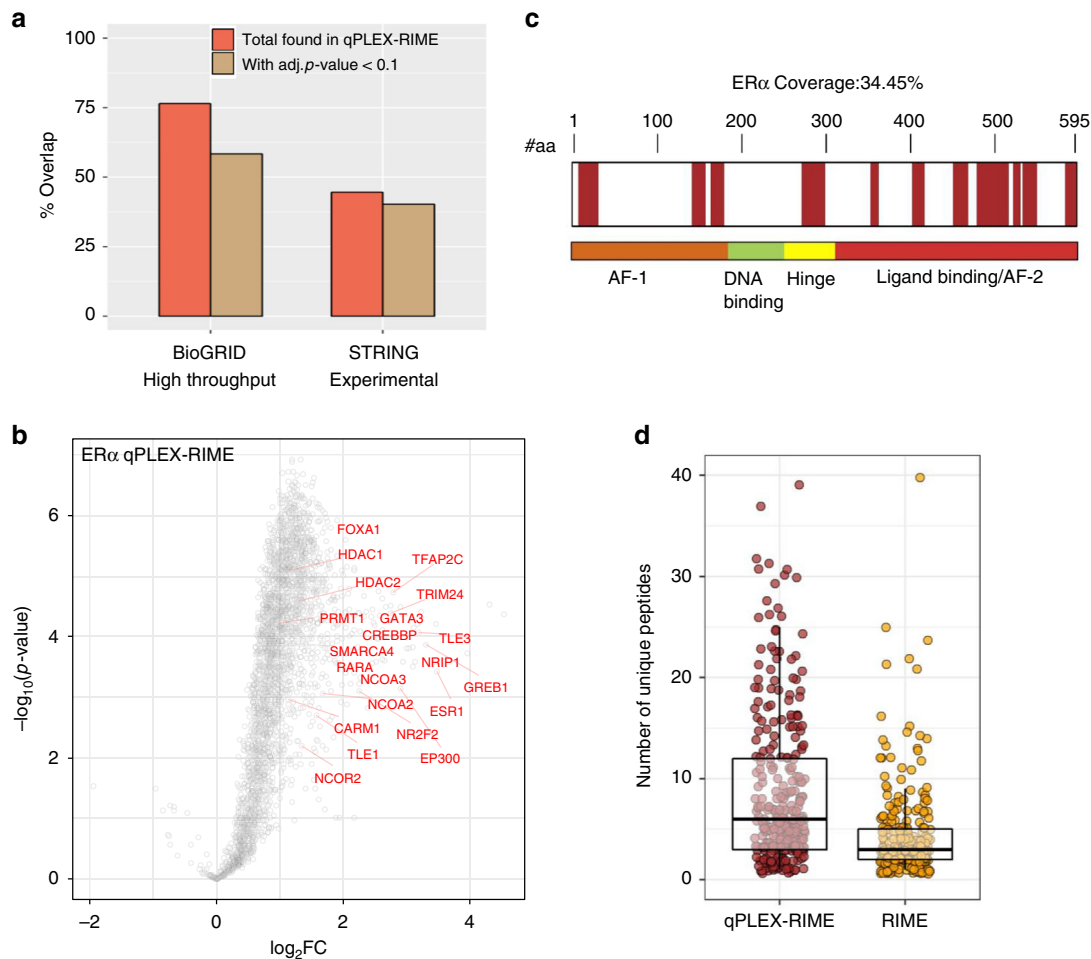
analysis (175 proteins >2-fold and adj.  $p$ -value < 0.01 (Limma moderated  $t$ -test) in qPLEX-RIME versus 51 proteins in non-quantitative RIME). Importantly, using the qPLEX-RIME we identified a number of novel ER $\alpha$ -associated candidate proteins. We validated the interactions of CBX3 (HP1 $\gamma$ ), NIPBL and FOXK1 with ER $\alpha$ , using Proximity Ligation Assay (PLA)<sup>29</sup> (Supplementary Fig. 1a). A GFP negative control was used to monitor for non-specific interactions (Supplementary Fig. 1b). Treatment of the MCF7 cells with the Selective ER $\alpha$  Degrader (SERD) Fulvestrant<sup>30</sup> (Supplementary Fig. 2) disrupted the above interactions demonstrating the specificity of the PLA assay and validating the interactors discovered by qPLEX-RIME (Supplementary Fig. 1a and c).

To test whether our quantitative pipeline can be widely used to study interactors of different bait proteins, we performed qPLEX-RIME experiments on three additional factors following the same experimental design as above. For these and all subsequent

experiments described in this study, we adopted the double crosslinking approach as a comparison between single and double crosslinking for ER $\alpha$  qPLEX-RIME data showed that the latter increases the pull-down efficiency of known and previously validated ER $\alpha$  interactors, including FOXA1, NR2F2 and NCOA2 (Supplementary Fig. 3a and Supplementary Data 3).

Firstly, the qPLEX-RIME method was applied to explore the interactome of CBP (CREB-binding protein) and NCOA3 (SRC-3); two well-characterised co-activators of nuclear receptors<sup>31</sup>. We identified 1437 and 1135 proteins for CBP and NCOA3, respectively, in the two multiplexed sets of bait and IgG pull-downs at peptide FDR < 1% (Supplementary Data 4 and 5). Both bait proteins were highly enriched in the target pull-downs compared to the IgG controls (CBP:log<sub>2</sub>fold-change = 3.2 and NCOA3:log<sub>2</sub>fold-change = 3.39) with a high number of unique peptides (44 unique peptides for CBP and 36 unique peptides for NCOA3) (Fig. 3a, b). Known interactors of CBP and NCOA3





**Fig. 2** Application of qPLEX-RIME for the identification of ER $\alpha$ -specific interactors in breast cancer cell lines. **a** Bar plots illustrating the overlap of the qPLEX-RIME data with known ER $\alpha$ -associated proteins from BioGRID and STRING databases. **b** Volcano plot summarising the quantitative results of the ER $\alpha$  qPLEX-RIME. ER $\alpha$  and several of its known significantly enriched interactors are labelled. **c** Sequence coverage of the ER $\alpha$  protein in the qPLEX-RIME analysis. **d** Boxplots of the number of unique peptides for the overlapping ER $\alpha$ -associated proteins between the non-quantitative RIME and the qPLEX-RIME method. Centre line shows the median, bounds of box correspond to the first and third quartiles and the upper and lower whiskers extend to the largest or the smallest value no further than  $1.5 \times \text{IQR}$  (inter-quartile range)

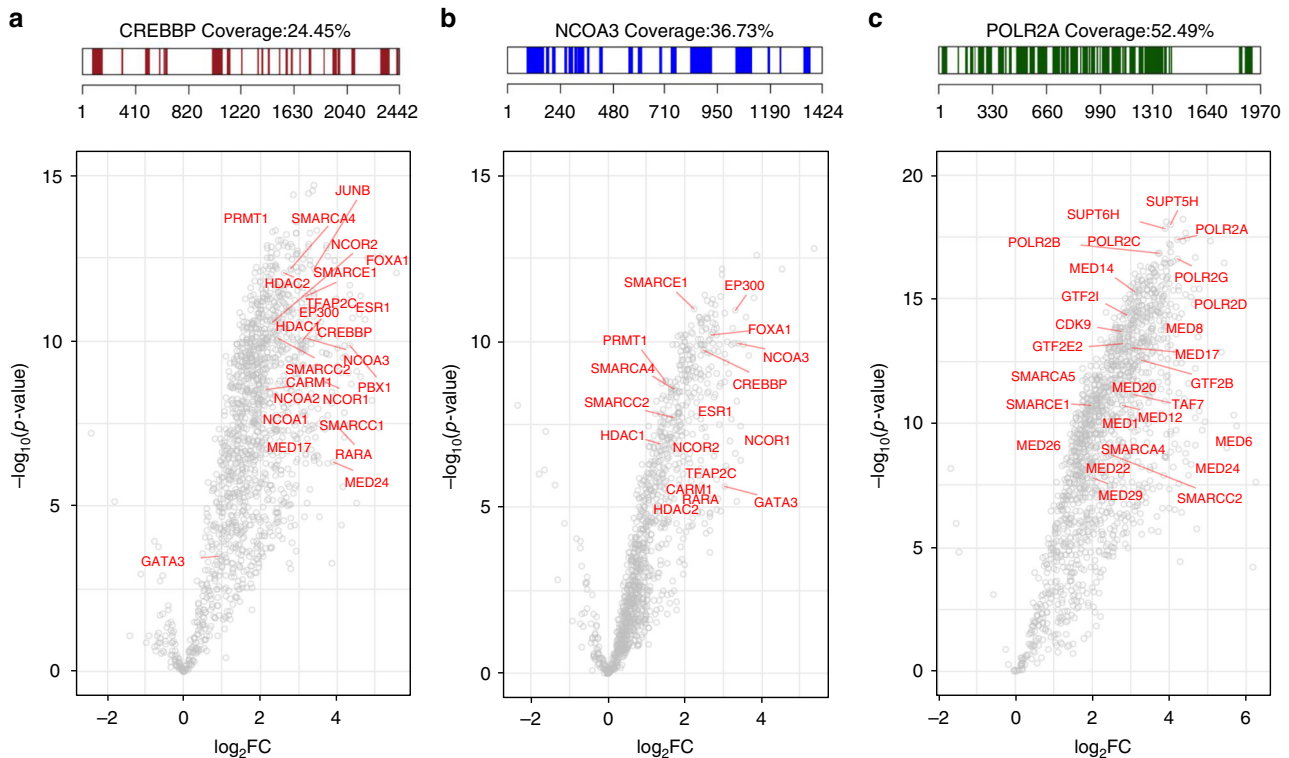
were identified including EP300, p160 co-activators, arginine methyltransferases and the ER $\alpha$  complex<sup>31</sup> (Fig. 3a, b). We also identified several members of the SWI/SNF chromatin remodelling complex, such as SMARCA4 (BRG1), SMARCE1 (BAF57), SMARCB1 (BAF47) and SMARCC2 (BAF170). Additionally, in the CBP qPLEX-RIME experiment we captured the association of CBP with the transcription factor JunB<sup>32</sup>, as well as with subunits of the mediator complex<sup>33</sup>, which are known to associate with enhancer regions as well. Interestingly, in addition to other co-activators, we also found a strong enrichment of co-repressors such as NCORs and HDACs in both data sets. This suggests that both co-activators and co-repressors are part of the same complex, which is consistent with previous findings demonstrating extensive co-localisation of co-repressors and co-activators by ChIP-seq<sup>34</sup>.

Secondly, we studied the interactome of phospho-RNA polymerase II (POLR2A) using an antibody that recognises the phosphorylated serine-5, which serves as a platform for assembly of factors that regulate transcription initiation, elongation, termination and mRNA processing<sup>35</sup>. We identified 1442 proteins across all multiplexed samples (Supplementary Data 6) and the bait protein was one of the top enriched proteins ( $\log_2\text{fold-change} = 4.2$ ), identified with 96 unique peptides

(Fig. 3c). A list of known polymerase II-associated factors were also observed, such as subunits of the SWI/SNF complex, proteins of the mediator complex, initiation and elongation factors<sup>36,37</sup> that are highlighted in Fig. 3c. A comparison of the interactomes of the four bait proteins (ER $\alpha$ , CBP, NCOA3 and POLR2A) showed significant numbers of uniquely identified interactors as well as partial overlap (Supplementary Fig. 3b). To examine whether the overlapping proteins are more likely due to the common underlying biology of the four baits rather than technical bias, we made a venn diagram using a random selection of proteins identified in the four qPLEX-RIME experiments, without considering enrichment relative to the IgG pull-downs (Supplementary Fig. 3c). This analysis showed a smaller number of proteins in the intersection of the four bait proteins indicating small contribution of technical factors to the observed overlap.

Taken together, our data demonstrate a gain in sensitivity using the qPLEX-RIME method that can lead to the identification of interacting proteins with statistical robustness and can be widely used for the characterisation of different interactomes.

**Study of ER $\alpha$  complex dynamics upon OHT treatment.** To investigate the dynamics of the ER $\alpha$  complex assembly upon



**Fig. 3** Application of qPLEX-RIME in CREBBP, NCOA3 and POLR2A. **a** Peptide sequence coverage of the CREBBP (CBP) protein in the qPLEX-RIME analysis (top panel). The volcano plot summarises the quantitative results of the CBP qPLEX-RIME and several of its known interactors are highlighted in red font (bottom panel). **b** Peptide sequence coverage of the NCOA3 protein in the qPLEX-RIME analysis (top panel). The volcano plot summarises the quantitative results of the NCOA3 qPLEX-RIME and several of its known interactors are highlighted in red font (bottom panel). **c** Peptide sequence coverage of the phospho-polymerase II (POLR2A) protein in the qPLEX-RIME analysis (top panel). The volcano plot summarises the quantitative results of the POLR2A qPLEX-RIME and several of its known interactors are highlighted in red font (bottom panel)

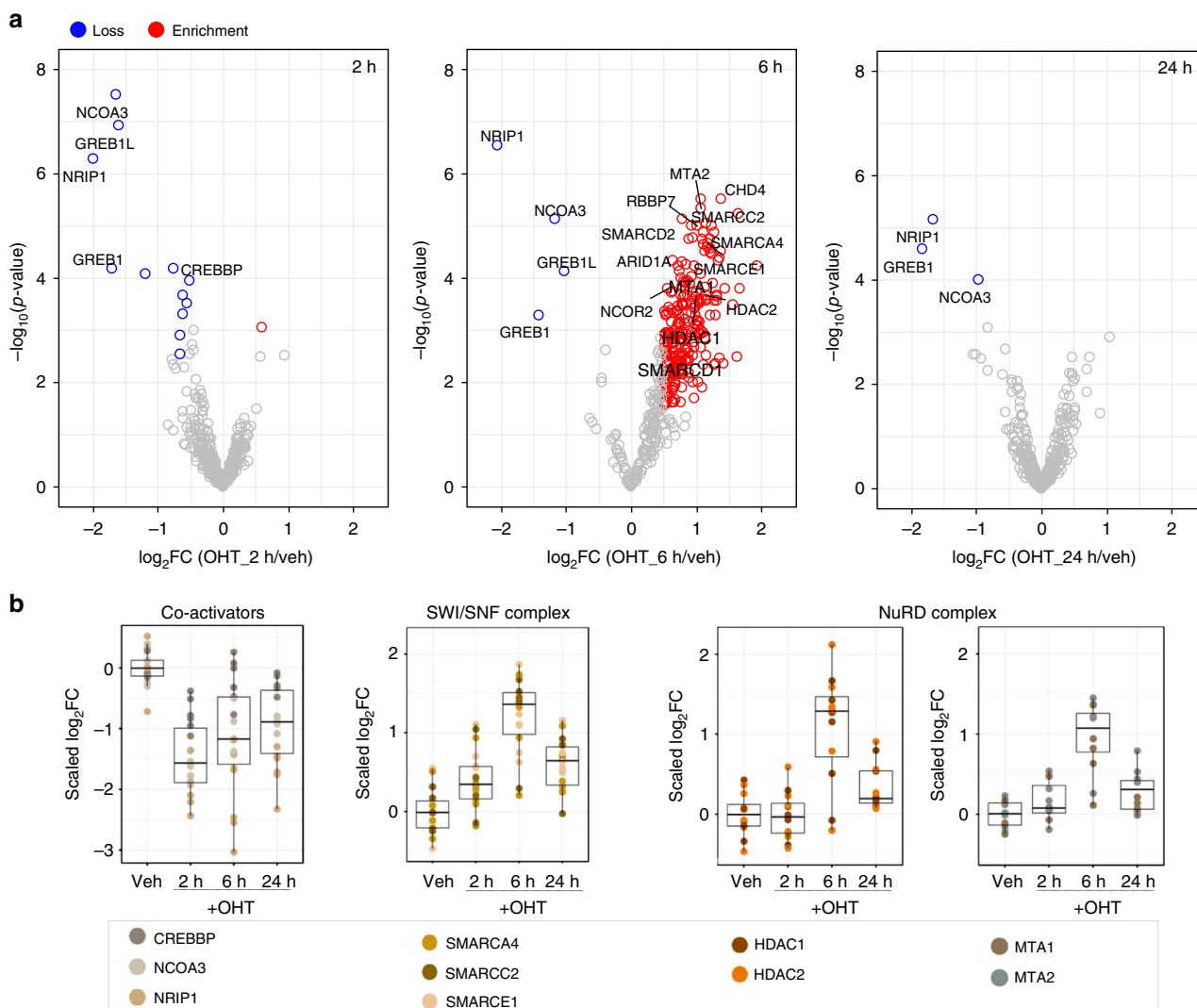
treatment with the Selective Estrogen Receptor Modulator (SERM) 4-hydroxytamoxifen (OHT), we performed three qPLEX-RIME experiments ( $3 \times 10$ plex) using independently prepared biological replicates. MCF7 cells were crosslinked after treatment with 100 nM OHT for 2 h, 6 h and 24 h or after 24 h of vehicle (ethanol) treatment. Two biological replicates of each condition were included in each experiment, resulting in a total of six replicates per time point. Additionally, MCF7 cells were treated with OHT or ethanol and crosslinked after 24 h treatment in each experiment to be used for control IgG pull-downs, to enable discrimination of non-specific binding.

To confirm that the drug treatment was successful, we performed RNA-seq analysis of six biological replicates using matched OHT treated samples. The mRNA data revealed transcriptional repression of a number of known ER $\alpha$  target genes at 6 h and 24 h, confirming the response to the drug treatment. Specifically, at 24 h treatment the expression of *PGR*, *PDZK1*, *TFF1*, *AREG*, *PKIB*, *SIAH2*, *MYB*, *HEY2*, *FOS*, *GREB1* and *TFF3*<sup>38–43</sup> was significantly inhibited compared to the vehicle treatment ( $\log_2$ Fold-Change  $< -0.5$ , adj.  $p$ -value  $< 0.05$ , Limma moderated  $t$ -test) (Supplementary Fig. 4a and Supplementary Data 7).

MultiNotch MS3 analysis of the qPLEX-RIME samples quantified 1105 proteins (FDR  $< 1\%$ ) across all three replicate experiments. Of these, 412 proteins were significantly enriched in ER $\alpha$  pull-downs compared to IgG samples ( $\log_2$ Fold-Change  $> 1$ , adj.  $p$ -value  $< 0.01$ , Limma moderated  $t$ -test) (Supplementary Data 8). Total ER $\alpha$  levels changed upon OHT treatment (Supplementary Fig. 4b), indicating that altered levels of antigen may influence the amount of purified proteins. Our data showed that this resulted in a significant dependency of the quantified

proteins on the amount of ER $\alpha$  pulled down (Supplementary Fig. 5a). To correct for this effect, we applied a linear regression approach<sup>15,44</sup> using the ER $\alpha$  profile as the independent variable and the profile of any other protein as the dependent variable. The advantage of this approach is that proteins with strong dependency on the target protein are subjected to significant correction, whereas proteins with small dependency on the target protein are only slightly corrected. Two such examples, of known ER $\alpha$  interactors before and after correction are shown in Supplementary Fig. 5b. Finally, using the quantification values corrected for the abundance of ER $\alpha$ , we found 249 specific proteins with altered profile in the interactome in at least one time point ( $|\log_2$ Fold-Change|  $> 0.5$ , adj.  $p$ -value  $< 0.05$ , Limma moderated  $t$ -test) allowing for a comprehensive mapping of the dynamic organisation of the ER $\alpha$  complex in response to OHT treatment.

**Dissociation and recruitment of co-factors upon OHT treatment.** We next interrogated the significant changes observed in the ER $\alpha$  interactome at each time point during OHT-mediated growth inhibition. After 2 h treatment with OHT, a significant loss of 12 proteins was observed including known ER $\alpha$  co-activators, such as NCOA3 (AIB1/SRC-3) and CREBBP (CBP) (Fig. 4a). These proteins have been associated with histone acetylation and activation of gene transcription<sup>23,45</sup> and their loss in the ER $\alpha$  interactome upon OHT treatment is consistent with previous studies showing that OHT binding blocks access of co-activators<sup>46</sup>. We also observed a significant loss of the interaction between ER $\alpha$  and NRIP1 (RIP140) protein. NRIP1 can act as a corepressor or as a coactivator<sup>47</sup> with previous evidence



**Fig. 4** Temporal profiling of the ER $\alpha$  interactome following treatment of MCF7 cells with OHT. **a** Volcano plots highlighting enriched or lost proteins in the ER $\alpha$  interactome upon OHT treatment for 2 h, 6 h and 24 h. **b** Boxplots illustrating the loss of ER $\alpha$  co-activator proteins CREBBP, NCOA3 and NRIP1 at 2 h and the enrichment of several subunits of the SWI/SNF and NuRD complexes at 6 h (left to right). Quantitative values are normalised so that the median of the vehicle treated samples is zero (centred around the median of vehicle). Centre line shows the median, bounds of box correspond to the first and third quartiles and the upper and lower whiskers extend to the largest or the smallest value no further than  $1.5 \times IQR$  (inter-quartile range)

suggesting that NRIP1 is required for ER $\alpha$ -complex formation and ER $\alpha$ -mediated gene expression<sup>47</sup>. The quantification profile of these key ER $\alpha$  co-activator proteins across all biological replicates is shown in Fig. 4b. Furthermore, OHT treatment resulted in the loss of GREB1 and its paralog gene product GREB1L. Loss of GREB1 upon OHT treatment has been previously described<sup>13</sup>, but here we report the loss of both proteins simultaneously.

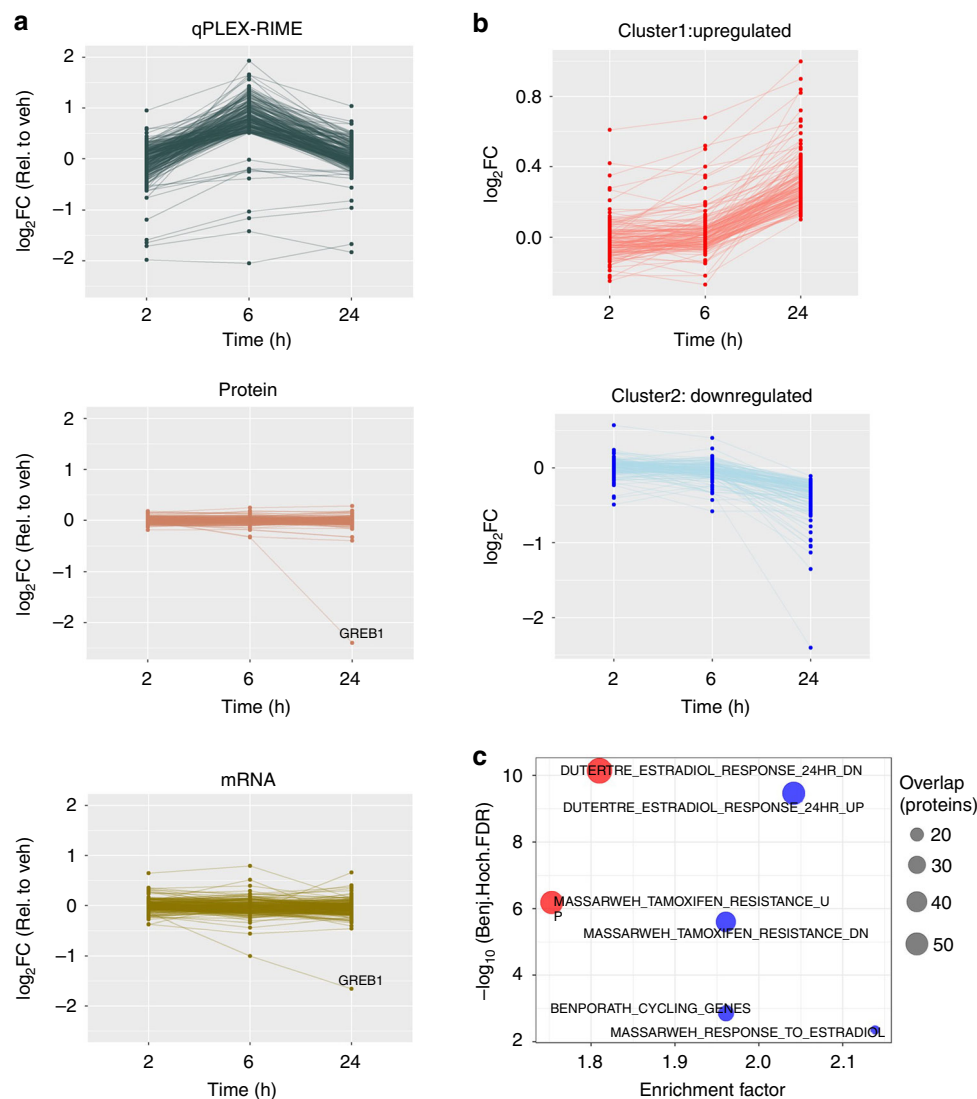
After 6 h treatment with OHT, 237 specific interactors showed significant enrichment ( $\log_2$ Fold-Change  $> 0.5$ , adj.  $p$ -value  $< 0.05$ , Limma moderated  $t$ -test) compared to the vehicle treatment, whilst NCOA3, NRIP1, GREB1 and GREB1L remained at decreased levels in the interactome ( $\log_2$ Fold-Change  $< -0.5$ , adj.  $p$ -value  $< 0.05$ , Limma moderated  $t$ -test) (Fig. 4a). Notably, there was an enrichment in the recruitment of several components of the NuRD (Nucleosome Remodelling and Deacetylase) complex, e.g. HDAC1/2<sup>24,48</sup> and the signature components MTA1/2<sup>24,48</sup>, as well as an enrichment of the co-repressor NCOR2 (SMRT)<sup>25,49</sup> (Fig. 4b). Consistently, NuRD complex and NCOR2 has been previously shown by ChIP to be

recruited to promoter regions of ER $\alpha$  target genes following OHT treatment<sup>23,24</sup>. Additionally, we found enriched subunits of the ATP-dependent chromatin remodelling complex SWI/SNF, which is known to regulate both gene activation and gene repression<sup>50,51</sup>. Detected components included SMARCC2 (BAF170), SMARCE1 (BAF57) and SMARCA4 (BRG1) (Fig. 4b). SMARCA4 protein, which was previously shown to be required for repression of ER-mediated transcription<sup>50</sup>, was one of the top enriched SWI/SNF proteins. The loss of NCOA3 and CBP at 2 h and the enrichment of SMARCC2 (BAF170) and HDAC1 at 6 h was validated with PLA assays (Supplementary Fig. 6a and b).

At 24 h we observed an almost complete restoration of the ER $\alpha$  complex, with the exception of the NCOA3, NRIP1 and GREB1 proteins, which were still decreased ( $\log_2$ Fold-Change  $< -0.5$ , adj.  $p$ -value  $< 0.05$ , Limma moderated  $t$ -test) (Fig. 4a). Taken together, our results indicate that the inhibitory effect of OHT peaks at 6 h, where ATP-dependent remodelling and corepressor complexes may coordinate to create a transcriptionally inactive chromatin environment.

**Identification of net changes in the ER $\alpha$  complex.** Our data suggest that treatment of MCF7 cells with OHT triggers significant changes in the composition of ER $\alpha$  interactome. To assess whether the changes identified by the qPLEX-RIME analysis are specific changes in interactions or result from changes in total protein levels, we performed timecourse whole proteome quantification in matched samples under the same conditions (vehicle, 2 h, 6 h and 24 h, four biological replicates each) (Supplementary Data 9). We confirmed the OHT up-regulation of ER $\alpha$  protein levels ( $\log_2$ Fold-Change: 2 h 0.21, 6 h 0.5 and 24 h 1), which was not due to an increase in gene transcription. This is consistent with previous reports demonstrating increased ER $\alpha$  stability in the presence of OHT<sup>52</sup>. A comparison between the qPLEX-RIME results and the total proteome data confirmed that the changes detected in the ER $\alpha$  complex upon OHT treatment represent changes in protein recruitment as the respective total protein and mRNA levels remained unchanged (Fig. 5a). GREB1 was the only ER $\alpha$  interactor with decreased mRNA and total protein levels at 24 h treatment. This is consistent with GREB1 being an ER $\alpha$  target gene<sup>13,43</sup> and explains the decreased association between ER $\alpha$  and GREB1 at this late time point.

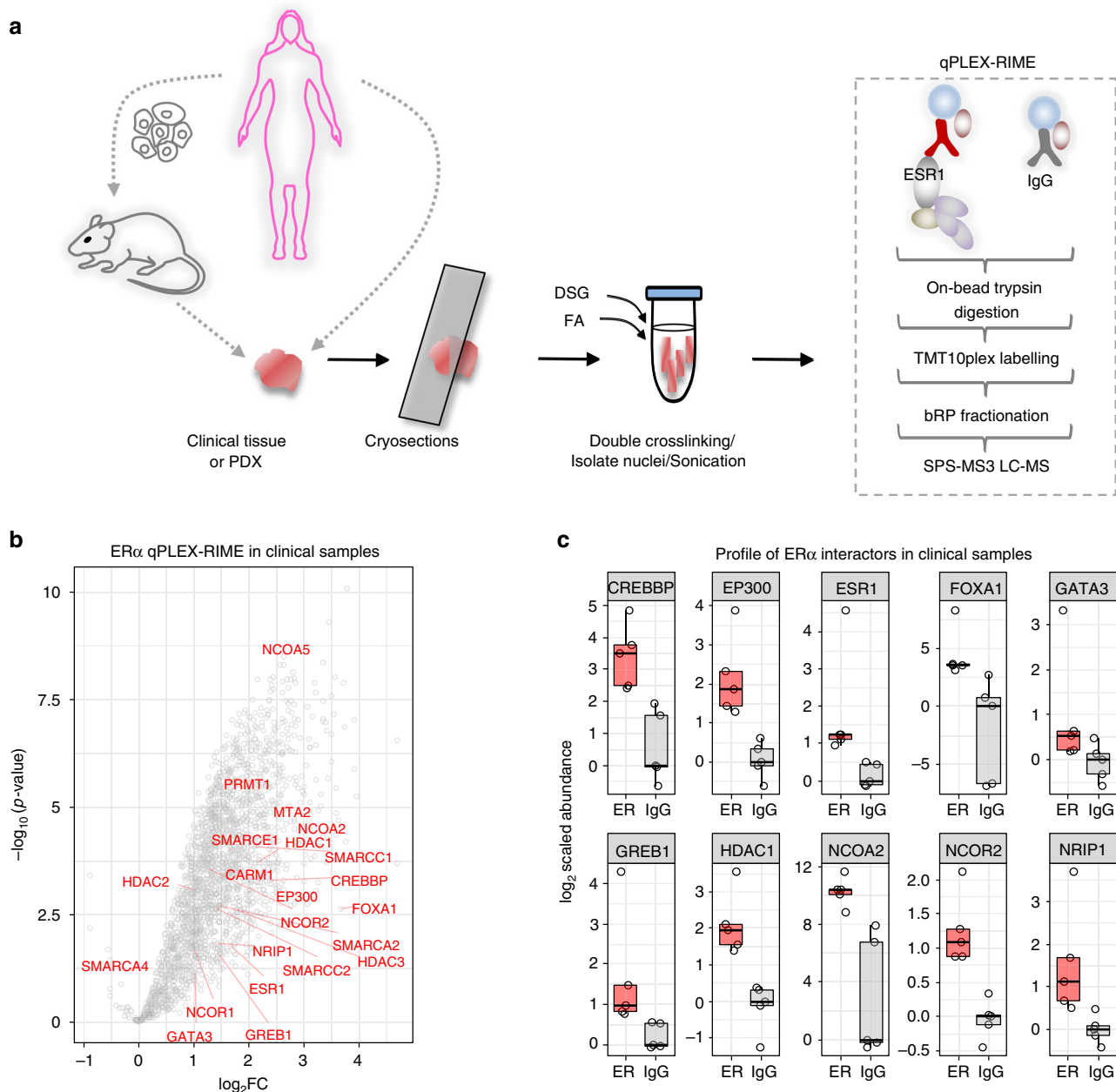
Downstream k-means clustering of the most variable proteins (adj.  $p$ -value < 0.05, Limma moderated  $t$ -test) across the three time points in the total proteome, identified clusters of up- and downregulated proteins (Fig. 5b). Gene Set Enrichment Analysis of the clusters, performed in Perseus software<sup>53</sup>, displayed an overrepresentation of genes related to estrogen response and tamoxifen resistance (Fig. 5c). Our findings also revealed the downregulation of proteins involved in cell cycle<sup>54</sup> (Supplementary Fig. 7a), in line with the antiproliferative effects of OHT<sup>24</sup>. Overall, significant changes in gene expression were observed already at 6 h coinciding with pronounced changes in the ER $\alpha$  interactome. As expected, the most significant changes in the total proteome were observed at the later time point (24 h). These results confirm that shuffling of ER $\alpha$ -associated proteins is not typically due to global changes in protein levels. The low mRNA-to-protein correlation at 2 h and 6 h and the respective strong correlation at 24 h are shown in Supplementary Fig. 7b and c. We conclude that our qPLEX-RIME data in combination with the total proteome measurements delineate both the local molecular events in the ER interactome and the associated downstream global effects.



**Fig. 5** Comparison of qPLEX-RIME data with total proteome analysis and RNA-seq data. **a** Line plots of the significantly enriched or lost proteins in the qPLEX-RIME data (top panel), their respective profiles in the total proteome analysis (middle panel) and in the RNA-seq analysis (bottom panel). **b** Line plots representing two k-means clusters of down- and upregulated proteins identified in total proteome analysis (top 10% most variable proteins with at least one significant change, adj.  $p$ -value < 0.05). **c** Gene set enrichment analysis for the down- and upregulated protein clusters

**Application of qPLEX-RIME in clinical tumour material.** To test whether qPLEX-RIME can be used to capture chromatin-associated protein–protein interactions in cancer specimens, we conducted an ER $\alpha$  qPLEX-RIME experiment using three independent ER positive human PDX tumours (HCI-003, HCI-005, HCI-006) that have been previously described<sup>55</sup>. Cryosections (30  $\mu$ m) of each tumour were double-crosslinked and each tumour was split into ER $\alpha$  and matched IgG pull-downs (Fig. 6a). The MultiNotch MS3 analysis identified 2319 proteins (FDR < 1%) across all multiplexed samples with highly reproducible profiles (Supplementary Fig. 8a and Supplementary Data 10). This analysis successfully recovered and quantified ER $\alpha$

(log<sub>2</sub>Fold-Change = 1.72, adj. *p*-value = 0.026, Limma moderated *t*-test, unique peptides = 3) using an unbiased mass spectrometry approach in tissue. In addition, many validated and known ER $\alpha$  interactors were discovered from the qPLEX-RIME conducted in PDX material, including CBP<sup>23,56</sup>, NCOA2<sup>56</sup>, HDAC1<sup>24</sup>, GREB1<sup>13</sup>, SMARCE1 (BAF57)<sup>51</sup>, SMARCA4 (BRG1)<sup>45</sup> and NCOA5 (CIA)<sup>57</sup> (Supplementary Fig. 8b). Sequence analysis of the qPLEX-RIME data showed that 60% of the significant interactors were identified with at least one unique human peptide (i.e. a peptide that does not align to the mouse proteome), indicating that the proteins identified above were primarily from the human cancer cells. Consistently, the tumour samples showed high



**Fig. 6** Characterisation of ER $\alpha$  interactors from in vivo samples. **a** Schematic representation of the ER $\alpha$  qPLEX-RIME workflow in human xenograft tissues or in human breast cancer tumours. **b** Volcano plot summarising the quantitative results of the ER $\alpha$  interactome in human breast cancer tumours. Several well-known ER $\alpha$  interactors are labelled. **c** Boxplots illustrating the enrichment of selected known ER $\alpha$  interactors in the ER $\alpha$  samples compared to IgG controls in human breast cancer tissues. The log<sub>2</sub> values are normalised so that the median of IgGs is zero. Centre line shows the median, bounds of box correspond to the first and third quartiles and the upper and lower whiskers extend to the largest or the smallest value no further than 1.5  $\times$  IQR (interquartile range)



cellularity and positive staining for human ER $\alpha$  exclusively in the cancer cells and not in the stroma (Supplementary Fig. 8c).

Prompted by the successful application of qPLEX-RIME in PDX tumours we sought to test the sensitivity of our method in human cancer clinical tissues, collected from surgery. To this end, we performed an ER $\alpha$  qPLEX-RIME experiment in five independent human breast cancer tumours (ER $\alpha$  positive, PR positive, Her2 negative and Grade2/Grade3). Approximately 60 sections (30  $\mu$ m) were obtained per sample, which were double-crosslinked and split for ER $\alpha$  and matched IgG pull-downs (Fig. 6a). The analysis successfully recovered ER $\alpha$  with excellent coverage (17 unique peptides), as well as 2191 proteins (FDR < 1%) that were quantified in all samples combined (Supplementary Data 11). These included well-described ER $\alpha$  interactors such as FOXA1, GATA3, GREB1, EP300, CBP, HDACs, NCOs and NCOA2 and subunits of the SWI/SNF complex (Fig. 6b). The enrichment of several ER $\alpha$  interactors in the bait samples compared to IgG control samples is illustrated in Fig. 6c.

Our data highlights the method's sensitivity and ability to identify endogenous protein networks from heterogeneous human tumour samples. Importantly, we report the identification of interactors from human tumour tissue material in an unbiased manner.

## Discussion

Here, we describe qPLEX-RIME, a proteomic method which enables comprehensive mapping of endogenous protein interactomes with high sensitivity and statistical robustness. The qPLEX-RIME approach integrates the well-established RIME immunoprecipitation method with advanced high-resolution quantitative multiplexed mass spectrometry analysis. The method can be utilised to discriminate enriched bona fide binding partners from contaminant proteins and to delineate the dynamics of chromatin-associated protein complexes with in-depth protein detection, reproducible quantification and increased sample throughput. The filtering criteria for the prioritisation of the best candidates depend on the type of experiment and the biological question. For bait proteins where very little is known about their interactome, we recommend the use of more stringent specificity criteria in terms of enrichment fold-change, *p*-value and number of unique peptides in combination with additional filtering based on functional annotations. When the focus is on the dynamic changes of interactomes, the prioritisation of the candidates mostly relies on their robust quantitative profiling across different conditions.

The multiplexed analysis of our pipeline eliminates the need to compare multiple data obtained by individual LC-MS runs, thereby increasing the quantification coverage in very low abundant protein interactors that are stochastically captured between independent replicate runs<sup>58</sup>. The ability to combine the labelled peptides derived from multiple samples increased the sensitivity of the method and enabled the characterisation of the ER $\alpha$  interactome in clinical tumours. Whilst interactors have previously been detected from clinical material, this required targeted mass spectrometry-based approaches and has not been done in an unbiased manner before<sup>13</sup>. Additionally, the use of isobaric labelling resolves the difficulties encountered with cell lines that are not compatible with stable-isotope labelled culture media and provides a means for quantitative analysis for clinical samples that are not amenable to in vivo isotopic labelling techniques. Importantly, our isobaric-labelling data demonstrated high reproducibility with previously published SILAC data<sup>13</sup> (Supplementary Fig. 7d), confirming the accurate quantification obtained by the MultiNotch MS3 level mass spectrometry analysis.

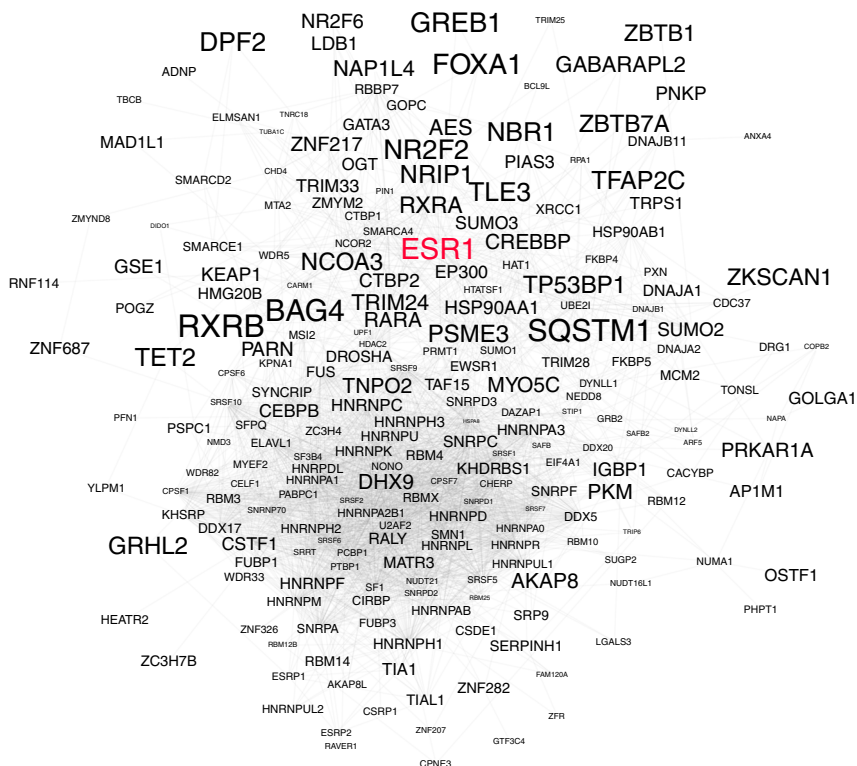
Here, we focused on ER $\alpha$ , the major driving transcription factor in luminal breast cancer<sup>56</sup>, which can be targeted by tamoxifen, a drug used for the treatment of ER<sup>+</sup> breast cancer<sup>59</sup>. Although many ER $\alpha$  interactors involved in ER-mediated gene expression have been discovered<sup>23,45</sup> our knowledge about their relevance at the tissue level and the impact of tamoxifen on their global association with ER $\alpha$  remains limited. The quantitative data obtained by the qPLEX-RIME experiments has provided us with a list of ER $\alpha$ -associated proteins with significant enrichment over the IgG samples. These include transient, indirect or weak interactions, as it is known that ER $\alpha$  associates with a number of different co-activators rapidly in a cyclic fashion<sup>23,45</sup>. As such the final readout of the crosslinking-based qPLEX-RIME method represents the sum of these interactions. Among these, we validated the interactions between ER $\alpha$  and three proteins; namely CBX3, NIPBL and FOXK1. CBX3 protein is a member of the HP1 protein family, a group of proteins that have been implicated in gene regulation, DNA replication and nuclear architecture<sup>60</sup>, whereas NIPBL is a core subunit of the highly conserved protein complex cohesin that has an important role in chromatin structure, gene expression, and DNA repair<sup>61</sup>. The transcription factor FOXK1 belongs to the forkhead family and has an important role in tumorigenesis<sup>62,63</sup>. These findings demonstrate that the gain in sensitivity obtained by qPLEX-RIME can reveal novel ER $\alpha$  interactors. Collectively, we identified a compendium of 253 proteins with consistent presence in all MCF7 data sets (Fig. 7). Importantly, our data show that the vast majority of these ER $\alpha$ -associated proteins (83%) can now be studied either in PDX or in human clinical tissues validating the relevance of these factors in vivo. Additionally, our qPLEX-RIME data on three additional factors, the CREBBP, NCOA3 and the phosphorylated form of POLR2A, highlight the wide applicability of our pipeline.

The application of qPLEX-RIME targeting ER $\alpha$  at multiple time points after OHT treatment, revealed a dynamic change in ER $\alpha$  co-regulators following drug treatment recapitulating and expanding the existing knowledge of OHT mechanism. After 2 h OHT treatment, we observed a loss of important transcriptional co-activators, such as NCOA3 and CBP, whereas at 6 h we observed enrichment on the recruitment of two well-conserved chromatin remodelling complexes, namely the NuRD and the SWI/SNF complex. This coincided with the enrichment of the basal corepressor NCOR2, which assists in the recruitment of HDAC proteins<sup>64</sup>. At the latest time point of 24 h, we observed a restoration of the ER $\alpha$  complex, which may be linked to the half-life of OHT. The exceptions were NRIP1, GREB1 and NCOA3. Interestingly, NCOA3 is amplified in breast cancer<sup>22</sup> and its expression levels have been associated with the effectiveness of tamoxifen treatment<sup>65</sup>. Further, ChIP-seq analysis has revealed that a number of binding sites of NCOA3 are associated with genes with a predictive value for breast cancer patient outcome<sup>56</sup>, supporting an important role of this co-regulator in tamoxifen response.

Our timecourse data indicate a switch between activation and repression of transcription in response to OHT treatment. This transition engages a two-step process with the immediate loss of co-activators, followed by the recruitment of co-repressors and ATP-chromatin remodelling complexes that may act cooperatively or in a sequential manner to accomplish transcriptional repression. The integration of qPLEX-RIME data with global protein and mRNA analysis provides a comprehensive view of the activity of a transcription-associated complex over time. A proposed model of OHT mechanism is depicted in Supplementary Fig. 9.

The qPLEX-RIME method can be used to monitor any dynamic changes of interest and importantly can be applied to clinical samples to study tumour evolution, treatment response or





**Fig. 7** Most frequently enriched ER $\alpha$  interactors. STRING network of 253 ER $\alpha$  interactors identified consistently across all the qPLEX-RIME analyses performed in MCF7 cells. The font size increases proportionally to the average fold-change enrichment of these proteins across all the ER $\alpha$  samples compare to IgG controls

numerous other biological and clinical questions. It provides a robust tool for the quantitative analysis of complexes that can be applied to generate comprehensive endogenous protein–protein interaction maps.

## Methods

**Cell lines and cell treatments.** ER $\alpha$ -expressing MCF7 breast cancer cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Gibco, Thermo Scientific, Leicestershire, UK, ref. 41965-239). Media was supplemented with 10% foetal bovine serum (FBS), 50 U/ml penicillin, 50  $\mu$ g/ml streptomycin and 2 mM L-glutamine. MCF7 cells were obtained from ATCC and they were tested for mycoplasma contamination. Also, the MCF7 cells were genotyped by short-tandem repeat genetic profiling using the PowerPlex\_16HS\_Cell Line panel and analysed using Applied Biosystems Gene Mapper ID v3.2.1 software by the external provider Genetica DNA Laboratories (LabCorp Specialty Testing Group). For the cell treatments, 4-Hydroxytamoxifen (Sigma-Aldrich, #HG278) or Fulvestrant ( Selleckchem, #S1191) were used at final concentration 100 nM.

**Whole cell lysate preparation and western blot analysis.** Cell pellets were reconstituted in 100  $\mu$ l RIPA buffer (Thermo Scientific, #89901) that was supplemented with protease inhibitors (Roche). 25  $\mu$ g protein from each sample was loaded on the gel (Invitrogen 4–12%) and the Precision Plus, Protein<sup>TM</sup> dual colour Standards Protein molecular weight marker (Bio-Rad, #161-0974) was used for the determination of protein sizes. The proteins were transferred onto a nitrocellulose membrane using the iBlot<sup>®</sup> 2 Dry Blotting System (Invitrogen) followed by one hour blocking using Odyssey<sup>®</sup> Blocking Buffer (Li-Cor, 927-40000). The membrane was immunoblotted with ER $\alpha$  antibody (Novocastra #6045332, 1:100) and beta-actin (Cell signalling #4970, 1:1000). Detection of the ER was achieved using the IRDye<sup>®</sup> 800 CW Goat anti-Mouse (926-32210, Li-Cor Biosciences) diluted to 1:5000, while the loading control was detected using the IRDye 680LT Goat anti-Rabbit (926-68071, Li-Cor Biosciences) diluted to 1:15000. All antibodies were diluted in Odyssey Buffer contained 0.1% Tween. Supplementary Fig. 10 shows the uncropped scan of the blot.

**RNA-seq analysis.** Cells were washed twice with cold Phosphate buffered saline (PBS) and harvested using 350  $\mu$ l of lysis buffer (RLT). Total RNA was extracted using the RNeasy<sup>®</sup> kit (Qiagen, #74106) according to the manufacturer's instructions. The extracted RNA was quantified using a NanoDrop<sup>®</sup> ND-1000

Spectrophotometer (Thermo Scientific). For the library preparation, the Illumina TruSeq Stranded mRNA Library Prep Kit High Throughput was used according to the manufacturer's instructions and two lanes of 50 bp single-end reads were run on HiSeq 4000. Reads were aligned to the human genome version GRCh37.75 using TopHat v2.1.0<sup>66</sup>. Read counts were obtained using feature Counts function in Subread v1.5.2<sup>67</sup> and read counts were normalised and tested for differential gene expression using the DESeq2 workflow<sup>68</sup>. Multiple testing correction was applied using the Benjamini–Hochberg method.

**RIME analysis.** MCF7 cells ( $2 \times 10^6$ ) were grown in complete media. The media was replaced with PBS containing 1% FA (Thermo #28908) and crosslinked for 10 min. For the double crosslinking cells were incubated in PBS containing 2 mM DSG (disuccinimidyl glutarate- Santa Cruz Biotechnology, #sc-285455A) for 20 min followed by incubation in 1% FA for 10 min. Crosslinking was quenched by adding glycine to a final concentration of 0.1 M. For the performance of RIME experiments, 50  $\mu$ l of Dynabeads<sup>®</sup> Protein A (Invitrogen) and 5  $\mu$ g of specific antibody were used for each sample. The antibodies used were: Rabbit polyclonal ER $\alpha$  antibody (Santa Cruz, sc-543), rabbit polyclonal SRC3 antibody (Bethyl laboratories, A300-347A), rabbit polyclonal CBP antibody (Diagenode, C15410224), rabbit polyclonal RNA polymerase II (phospho S5) antibody (Abcam, ab5131) and rabbit IgG antibody (Santa Cruz, sc-2027 or Abcam, ab171870). For nuclear extraction the cell pellet was resuspended in LB1 buffer (50 mM HEPES-KOH (pH 7.5), 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40 and 0.25% Triton X-100) followed by rotation mixing for 10 min at 4 °C. Then, nuclei were pelleted and resuspended in LB2 buffer (10 mM Tris-HCL (pH 8.0), 200 mM NaCl, 1 mM EDTA and 0.5 mM EGTA) and rotated at 4 °C for 5 min. The samples were resuspended in LB3 buffer (10 mM Tris-HCL (pH 8), 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-deoxycholate and 0.5% N-lauroylsarcosine). Chromatin was sheared by sonication (Diagenode) to produce DNA fragments of 100–1,000 bp. The bead-bound antibody and chromatin were incubated overnight at 4 °C. The next day the beads were washed 10 times with 1 ml ice-cold RIPA buffer and twice with 500  $\mu$ l 100 mM AMBIC (ammonium bicarbonate).

**Proximity ligation assay.** Cells were fixed and permeabilised by the addition of ice-cold methanol (–20 °C) for 3 min followed by three washing steps with cold PBS. PLA was carried out according to manufacturer's instructions (Sigma Aldrich, #DUO92007). The following primary antibodies were used for the PLA assay: ER $\alpha$  (Santa Cruz, sc-543 or sc-8002, 1:250) HP1 $\gamma$  (Santa Cruz, sc-365085, 1:400), NIPBL (Santa Cruz, sc-374625, 1:200), FOXK1 (Santa Cruz, sc-373810, 1:200), GFP

(Abcam, ab1218, 1:200), NCOA3 (Bethyl Laboratories, A300-347A, 1:200), CBP (Bethyl Laboratories, A300-363A, 1:200), BAF170 (Santa Cruz, sc-17838, 1:200), HDAC1 (Santa Cruz, sc-81598, 1:200) and the incubation on the coverslips was performed for 1 h at 37 °C. For the single PLA recognition experiment two ER $\alpha$  antibodies (Santa Cruz, sc-543, 1:800 and Invitrogen, MA5-13191, 1:1200) were used in combination. The secondary proximity probes (Sigma Aldrich, Rabbit-PLUS, #DUO92002 and Mouse-MINUS, #DUO92004) were incubated for 1 h at 37 °C. The Leica DFC340FX microscope was used and images were captured with Leica Imaging software. DAPI and PLA fluorescence were captured at high resolution for a total of 8 separate observation fields. Cell numeration and PLA labelling were carried out using Image J software. Cells and red PLA dots were counted using the 'Analyze Particles' function. For each condition at least 200 cells were imaged and analysed. Then, the average value of number of spots per nucleus was calculated. All statistical analyses were carried out by performing Student's *t*-test.

**Immunofluorescence.** Cells were fixed and permeabilised with ice-cold (−20 °C) methanol for 3 min and after fixation cells were blocked in PBS-5% (w/v) Bovine Serum Albumin (BSA) (Sigma-Aldrich) for 30 min at room temperature. The primary ER $\alpha$  antibody (Santa Cruz, sc-543, 1:250) was diluted in blocking solution (PBS-5% (w/v) BSA) and incubated on coverslips for 1 h at 37 °C. Afterwards, the coverslips were washed four times in washing buffer (PBS-0.5% Tween). Secondary antibody conjugated to Alexa Fluor 488 (Invitrogen, #A-21206, 1:500) was diluted in blocking solution and incubated on coverslips for 1 h at 37 °C in the dark. Coverslips were then washed again three times in washing buffer and once in PBS.

**PDX propagation and tissue collection.** Viable frozen PDX tumour tissue was propagated in immune-compromised mice. Briefly, 1 mm<sup>3</sup> tumour pieces were implanted into the 4<sup>th</sup> mammary pad of NSG mice. All mice were supplemented with estrogen, using silastic E2 pellets (made in-house) inserted into the dorsal scruff. Twice weekly standard monitoring and tumour measurement was conducted. Once tumours reached appropriate size, ~1000 mm<sup>3</sup>, mice were sacrificed by cervical dislocation under deep, isoflurane-induced anaesthesia. The tumours were resected, diced and processed by either snap freezing in liquid nitrogen, fixing in 10% neutral buffered formalin solution for subsequent paraffin embedding, embedding in OCT, or viably freezing in FCS supplemented with 5% DMSO.

**Sample preparation of clinical tumour material.** Clinical samples were cryosectioned in 30  $\mu$ m slices using the Leica CM 3050 S cryostat. Tissue sections were fixed in a two-step procedure by adding 2 mM DSG for 25 min. In the same suspension of tissue sections, 1% FA was added for another 20 min without removal of the DSG. Crosslinking was quenched by the addition of glycine to a final concentration of 0.25 M. Samples were centrifuged for 3 min at 2500 g and the supernatant was discarded. Tissue pellets were washed twice with cold PBS and resuspended in 6 ml LB3 buffer (10 mM Tris-HCl (pH 8), 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-deoxycholate, and 0.5% N-lauroylsarcosine), followed by tip sonication for 12–20 cycles (30 s on, 1 min off) depending on the tumour size. The downstream processing was performed as described above (see RIME method section) and the tissue samples were separated in two parts for the performance of ER $\alpha$  and IgG RIME pull-down assays. Patient and patient-derived tissues used in this work were collected under protocol X13-0133, HREC/13/RPAH/187. HREC approval was obtained through the SLHD (Sydney Local Health District) Ethics Committee (Royal Prince Alfred Hospital) zone, and site-specific approvals were obtained for all additional sites. Written consent was obtained from all patients prior to collection of tissue and clinical data stored in a de-identified manner, following pre-approved protocols. All animal procedures were carried out in accordance to relevant national and international guidelines and animal protocols approved by the Garvan/St Vincent's Animal Ethics Committee (Animal ethics number 15/10).

**Immunohistochemistry.** FFPE blocks from PDX tumours were sectioned at 4  $\mu$ m onto Superfrost Plus slides. Immunohistochemistry was carried out using the Leica Bond Autostainer. Sections underwent dewaxing, heat induced antigen retrieval (Leica reagent ER2, 30 mins), and primary and secondary antibody incubations, using ER $\alpha$  antibody (ab108398, Abcam, 1:500) and the EnVision + Rabbit secondary system, respectively. Sections were counterstained with haematoxylin.

**Trypsin digestion and TMT labelling.** A volume of 10  $\mu$ l trypsin solution (15 ng/ $\mu$ l) (Pierce) in 100 mM AMBIC was added to the beads followed by overnight incubation at 37 °C. A second digestion step was performed the next day for 4 h. After proteolysis the tubes were placed on a magnet and the supernatant solution was collected after acidification by the addition of 2  $\mu$ l 5% formic acid. The resultant peptides were cleaned with the Ultra-Micro C18 Spin Columns (Harvard Apparatus) according to manufacturer's instructions. The peptide samples were dried with speedvac, reconstituted in 100  $\mu$ l 0.1 M TEAB (triethylammonium bicarbonate) and labelled using the TMT-10plex reagents (Thermo Fisher) with a randomised design. The peptide mixture was fractionated with Reversed-Phase cartridges at high pH (Pierce #84868). Nine fractions were collected using different elution solutions in the range of 5–50% ACN.

For the total proteome analysis 200  $\mu$ l of 0.1 M TEAB, 0.1% SDS buffer was added to each cell pellet followed by probe sonication and boiling at 95 °C. Protein concentration was estimated with Bradford assay (BIO-RAD-Quick start) according to manufacturer's instructions. For each sample, 90  $\mu$ g of total protein were reduced for 1 h at 60 °C by the addition of 2  $\mu$ l 50 mM tris-2-carboxyethyl phosphine (TCEP, Sigma). Cysteines were blocked for 10 min on the bench with the addition of 1  $\mu$ l 200 mM methyl methanethiosulfonate (MMTS, Sigma). For peptide generation, trypsin (Pierce #90058) solution was added at ratio protein/trypsin ~30:1 for overnight digestion at 37 °C. The next day peptides were allowed to react with the TMT-10plex reagents (Thermo Scientific) for one hour. The reaction was quenched with 8  $\mu$ l of 5% hydroxylamine (Thermo Scientific) and the labelled samples were mixed and dried with speedvac concentrator. The TMT mix was reconstituted and fractionated on a Dionex Ultimate 3000 system at high pH using the X-Bridge C18 column (3.5  $\mu$ m 2.1  $\times$  150 mm, Waters) with 1% gradient. UV signal was recorded at 280 and 215 nm and fractions were collected in a peak dependent manner.

**LC-MS analysis.** Peptide fractions were analysed on a Dionex Ultimate 3000 UHPLC system coupled with the nano-ESI Fusion Lumos (Thermo Scientific). Samples were loaded on the Acclaim PepMap 100, 100  $\mu$ m  $\times$  2 cm C18, 5  $\mu$ m, 100  $\text{\AA}$  trapping column with the ulPickUp injection method using the loading pump at 5  $\mu$ l/min flow rate for 10 min. For the peptide separation the EASY-Spray analytical column 75  $\mu$ m  $\times$  25 cm, C18, 2  $\mu$ m, 100  $\text{\AA}$  column was used for multi-step gradient elution. Mobile phase (A) was composed of 2% acetonitrile, 0.1% formic acid and mobile phase (B) was composed of 80% acetonitrile, 0.1% formic acid. The elution method at flow rate 300 nL/min included the following: for 95 min gradient up to 45% (B), for 5 min gradient up to 95% (B), for 8 min isocratic 95% (B), for 2 min down to 5% (B), for 10 min isocratic equilibration 5% (B) at 40 °C. For the clinical sample analysis, a longer gradient separation was used as follows: for 160 min gradient up to 40% (B), for 10 min gradient up to 95% (B), for 8 min isocratic 95% (B), for 2 min down to 5% (B), and for 10 min isocratic equilibration 5% (B). The Lumos was operated in a data-dependent mode for both MS2 and SPS-MS3 methods. The full scans were performed in the Orbitrap in the range of 380–1500 *m/z* at 120 K resolution. The MS2 scans were performed in the ion trap with collision energy 35%. Peptides were isolated in the quadrupole with isolation window 0.7 Th. The 10 most intense fragments were selected for Synchronous Precursor Selection (SPS) HCD-MS3 analysis with MS2 isolation window 2.0 Th. The HCD collision energy was set at 55% and the detection was performed with Orbitrap resolution 60k and in scan range 110–400 *m/z*.

**Data processing and interpretation.** The collected CID tandem mass spectra were processed with the SequestHT search engine on the Proteome Discoverer 2.1 software for peptide and protein identifications. The node for SequestHT included the following parameters: Precursor Mass Tolerance 20 ppm, Fragment Mass Tolerance 0.5 Da, Dynamic Modifications were Oxidation of M (+15.995 Da), Deamidation of N, Q (+0.984 Da) and Static Modifications were TMT6plex at any N-Terminus, K (+229.163 Da) for the quantitative data. Methylthio at C (+45.988) was included for the total proteome data. The Reporter Ion Quantifier node included a TMT 6plex (Thermo Scientific Instruments) Quantification Method, for MS3 scan events, HCD activation type, integration window tolerance 20 ppm and integration method Most Confident Centroid. The consensus workflow included S/N calculation for TMT intensities and the level of confidence for peptide identifications was estimated using the Percolator node with decoy database search. Strict FDR was set at *q*-value < 0.01.

**Bioinformatics Analysis.** We developed an R package (qPLEXanalyzer) to perform downstream data analysis. All analyses were performed using only unique peptides identified with high confidence (peptide FDR < 1%) across all experiments. Peptide-level signal-to-noise (S/N) TMT values were corrected for equal loading across samples using different normalisation approaches based upon the experiment type. For the regression-based correction, unique peptides were aggregated and proteins identified in all the experiments were kept for further analysis. The normalisation on the bait protein level was carried out at protein level using log<sub>2</sub> row-mean scaled values. To filter-out non-specific proteins, a limma-based differential analysis was performed comparing ER and IgG control samples. In the regression analysis, the ER $\alpha$  profile was used as the independent variable (*x*) and the profile of any other protein as the dependent variable (*y*) excluding the IgG controls. The residuals of the  $y = ax + b$  linear model represent the protein quantification profiles that are not driven by ER $\alpha$  amount in the pull-down. More details on the normalisation methods used can be found in Supplementary Note 2. The identification of differentially bound proteins was carried out using the limma-based analysis. A multiple testing correction was applied on *p*-value using the Benjamini–Hochberg method to control the FDR.

**Code availability.** The qPLEXanalyzer and qPLEXdata R packages are available at (<https://doi.org/10.5281/zenodo.1237825>). Both pipelines are described in detail in Supplementary Notes 1 and 2.

**Data availability.** RNA-seq data have been deposited in NCBI's Gene Expression Omnibus<sup>69</sup> and are accessible through GEO Series accession number [GSE104872](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE104872). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE<sup>70</sup> partner repository with the data set identifier [PXD007968](https://www.ebi.ac.uk/pride/archive/projects/PXD007968). All other data supporting the findings of this study are available from the corresponding authors on reasonable request.

Received: 2 November 2017 Accepted: 3 May 2018

Published online: 13 June 2018

## References

- Vidal, M., Cusick, M. E. & Barabasi, A. L. Interactome networks and human disease. *Cell* **144**, 986–998 (2011).
- Ewing, R. M. et al. Large-scale mapping of human protein–protein interactions by mass spectrometry. *Mol. Syst. Biol.* **3**, 89 (2007).
- Stelzl, U. et al. A human protein–protein interaction network: a resource for annotating the proteome. *Cell* **122**, 957–968 (2005).
- Rolland, T. et al. A proteome-scale map of the human interactome network. *Cell* **159**, 1212–1226 (2014).
- Rual, J. F. et al. Towards a proteome-scale map of the human protein–protein interaction network. *Nature* **437**, 1173–1178 (2005).
- Hein, M. Y. et al. A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell* **163**, 712–723 (2015).
- Huttlin, E. L. et al. Architecture of the human interactome defines protein communities and disease networks. *Nature* **545**, 505–509 (2017).
- Collins, B. C. et al. Quantifying protein interaction dynamics by SWATH mass spectrometry: application to the 14-3-3 system. *Nat. Methods* **10**, 1246–1253 (2013).
- Liu, F., Rijkers, D. T., Post, H. & Heck, A. J. Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. *Nat. Methods* **12**, 1179–1184 (2015).
- Rafiee, M. R., Girardot, C., Sigismondo, G. & Krijgsvelde, J. Expanding the circuitry of pluripotency by selective isolation of chromatin-associated proteins. *Mol. Cell* **64**, 624–635 (2016).
- Wierer, M. & Mann, M. Proteomics to study DNA-bound and chromatin-associated gene regulatory complexes. *Hum. Mol. Genet.* **25**, R106–R114 (2016).
- Engelen, E. et al. Proteins that bind regulatory regions identified by histone modification chromatin immunoprecipitations and mass spectrometry. *Nat. Commun.* **6**, 7155 (2015).
- Mohammed, H. et al. Endogenous purification reveals GREB1 as a key estrogen receptor regulatory factor. *Cell Rep.* **3**, 342–349 (2013).
- Mohammed, H. et al. Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. *Nat. Protoc.* **11**, 316–326 (2016).
- Roumeliotis, T. I. et al. Genomic determinants of protein abundance variation in colorectal cancer cells. *Cell Rep.* **20**, 2201–2214 (2017).
- McAlister, G. C. et al. Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. *Anal. Chem.* **84**, 7469–7478 (2012).
- McAlister, G. C. et al. MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. *Anal. Chem.* **86**, 7150–7158 (2014).
- Nowak, D. E., Tian, B. & Brasier, A. R. Two-step cross-linking method for identification of NF-kappaB gene network by chromatin immunoprecipitation. *Biotechniques* **39**, 715–725 (2005).
- Chatr-Aryamontri, A. et al. The BioGRID interaction database: 2017 update. *Nucleic Acids Res.* **45**, D369–D379 (2017).
- Szklarczyk, D. et al. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Res.* **45**, D362–D368 (2017).
- Tsai, W. W. et al. TRIM24 links a non-canonical histone signature to breast cancer. *Nature* **468**, 927–932 (2010).
- Anzick, S. L. et al. AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer. *Science* **277**, 965–968 (1997).
- Shang, Y., Hu, X., DiRenzo, J., Lazar, M. A. & Brown, M. Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription. *Cell* **103**, 843–852 (2000).
- Liu, X. F. & Bagchi, M. K. Recruitment of distinct chromatin-modifying complexes by tamoxifen-complexed estrogen receptor at natural target gene promoters in vivo. *J. Biol. Chem.* **279**, 15050–15058 (2004).
- Jepsen, K. et al. Combinatorial roles of the nuclear receptor corepressor in transcription and development. *Cell* **102**, 753–763 (2000).
- Carroll, J. S. et al. Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell* **122**, 33–43 (2005).
- Tan, S. K. et al. AP-2gamma regulates oestrogen receptor-mediated long-range chromatin interaction and gene transcription. *EMBO J.* **30**, 2569–2581 (2011).
- Theodorou, V., Stark, R., Menon, S. & Carroll, J. S. GATA3 acts upstream of FOXA1 in mediating ESR1 binding by shaping enhancer accessibility. *Genome Res.* **23**, 12–22 (2013).
- Soderberg, O. et al. Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nat. Methods* **3**, 995–1000 (2006).
- Wardell, S. E., Marks, J. R. & McDonnell, D. P. The turnover of estrogen receptor alpha by the selective estrogen receptor degrader (SERD) fulvestrant is a saturable process that is not required for antagonist efficacy. *Biochem. Pharmacol.* **82**, 122–130 (2011).
- Lonard, D. M. & O'Malley, B. W. Nuclear receptor coregulators: judges, juries, and executioners of cellular regulation. *Mol. Cell* **27**, 691–700 (2007).
- Bannister, A. J., Oehler, T., Wilhelm, D., Angel, P. & Kouzarides, T. Stimulation of c-Jun activity by CBP: c-Jun residues Ser63/73 are required for CBP induced stimulation in vivo and CBP binding in vitro. *Oncogene* **11**, 2509–2514 (1995).
- Acevedo, M. L. & Kraus, W. L. Mediator and p300/CBP-steroid receptor coactivator complexes have distinct roles, but function synergistically, during estrogen receptor alpha-dependent transcription with chromatin templates. *Mol. Cell Biol.* **23**, 335–348 (2003).
- Siersbaek, R. et al. Dynamic rewiring of promoter-anchored chromatin loops during adipocyte differentiation. *Mol. Cell* **66**, 420–435 e425 (2017).
- Phatnani, H. P. & Greenleaf, A. L. Phosphorylation and functions of the RNA polymerase II CTD. *Genes & Dev.* **20**, 2922–2936 (2006).
- Orphanides, G. & Reinberg, D. RNA polymerase II elongation through chromatin. *Nature* **407**, 471–475 (2000).
- Hahn, S. Structure and mechanism of the RNA polymerase II transcription machinery. *Nat. Struct. Mol. Biol.* **11**, 394–403 (2004).
- Frasor, J. et al. Selective estrogen receptor modulators: discrimination of agonistic versus antagonistic activities by gene expression profiling in breast cancer cells. *Cancer Res.* **64**, 1522–1533 (2004).
- Frasor, J. et al. Profiling of estrogen up- and down-regulated gene expression in human breast cancer cells: insights into gene networks and pathways underlying estrogenic control of proliferation and cell phenotype. *Endocrinology* **144**, 4562–4574 (2003).
- Frasor, J., Danes, J. M., Funk, C. C. & Katzenellenbogen, B. S. Estrogen down-regulation of the corepressor N-CoR: mechanism and implications for estrogen derepression of N-CoR-regulated genes. *Proc Natl Acad. Sci. USA* **102**, 13153–13157 (2005).
- Caliceti, C. et al. 17beta-estradiol enhances signalling mediated by VEGF-A-delta-like ligand 4-notch1 axis in human endothelial cells. *PLoS ONE* **8**, e71440 (2013).
- May, F. E. & Westley, B. R. Expression of human intestinal trefoil factor in malignant cells and its regulation by oestrogen in breast cancer cells. *J. Pathol.* **182**, 404–413 (1997).
- Ghosh, M. G., Thompson, D. A. & Weigel, R. J. PDZK1 and GREB1 are estrogen-regulated genes expressed in hormone-responsive breast cancer. *Cancer Res.* **60**, 6367–6375 (2000).
- McCarthy, D. J., Campbell, K. R., Lun, A. T. & Wills, Q. F. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* **33**, 1179–1186 (2017).
- Metivier, R. et al. Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. *Cell* **115**, 751–763 (2003).
- Shiau, A. K. et al. The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* **95**, 927–937 (1998).
- Rosell, M. et al. Complex formation and function of estrogen receptor alpha in transcription requires RIP140. *Cancer Res.* **74**, 5469–5479 (2014).
- Lai, A. Y. & Wade, P. A. Cancer biology and NuRD: a multifaceted chromatin remodeling complex. *Nat. Rev. Cancer* **11**, 588–596 (2011).
- Chen, J. D. & Evans, R. M. A transcriptional co-repressor that interacts with nuclear hormone receptors. *Nature* **377**, 454 (1995).
- Zhang, B., Chambers, K. J., Faller, D. V. & Wang, S. Reprogramming of the SWI/SNF complex for co-activation or co-repression in prohibitin-mediated estrogen receptor regulation. *Oncogene* **26**, 7153–7157 (2007).
- Belandia, B., Orford, R. L., Hurst, H. C. & Parker, M. G. Targeting of SWI/SNF chromatin remodelling complexes to estrogen-responsive genes. *EMBO J.* **21**, 4094–4103 (2002).
- Wijayaratne, A. L. & McDonnell, D. P. The human estrogen receptor-alpha is a ubiquitinated protein whose stability is affected differentially by agonists,



- antagonists, and selective estrogen receptor modulators. *J. Biol. Chem.* **276**, 35684–35692 (2001).
53. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
  54. Whitfield, M. L. et al. Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol. Biol. Cell.* **13**, 1977–2000 (2002).
  55. DeRose, Y. S. et al. Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes. *Nat. Med.* **17**, 1514–1520 (2011).
  56. Zwart, W. et al. Oestrogen receptor-co-factor-chromatin specificity in the transcriptional regulation of breast cancer. *EMBO J.* **30**, 4764–4776 (2011).
  57. Sauve, F. et al. CIA, a novel estrogen receptor coactivator with a bifunctional nuclear receptor interacting determinant. *Mol. Cell. Biol.* **21**, 343–353 (2001).
  58. Rauniyar, N. & Yates, J. R. 3rd Isobaric labeling-based relative quantification in shotgun proteomics. *J. Proteome Res.* **13**, 5293–5309 (2014).
  59. Shang, Y. & Brown, M. Molecular determinants for the tissue specificity of SERMs. *Science* **295**, 2465–2468 (2002).
  60. Eissenberg, J. C. & Elgin, S. C. The HP1 protein family: getting a grip on chromatin. *Curr. Opin. Genet. Dev.* **10**, 204–210 (2000).
  61. Hill, V. K., Kim, J. S. & Waldman, T. Cohesin mutations in human cancer. *Biochim. Biophys. Acta* **1866**, 1–11 (2016).
  62. Katoh, M. & Katoh, M. Identification and characterization of human FOXX1 gene in silico. *Int. J. Mol. Med.* **14**, 127–132 (2004).
  63. Wu, Y. et al. Oncogene FOXX1 enhances invasion of colorectal carcinoma by inducing epithelial-mesenchymal transition. *Oncotarget* **7**, 51150–51162 (2016).
  64. Underhill, C., Qutob, M. S., Yee, S. P. & Torchia, J. A novel nuclear receptor corepressor complex, N-CoR, contains components of the mammalian SWI/SNF complex and the corepressor KAP-1. *J. Biol. Chem.* **275**, 40463–40470 (2000).
  65. Osborne, C. K. et al. Role of the estrogen receptor coactivator AIB1 (SRC-3) and HER-2/neu in tamoxifen resistance in breast cancer. *J. Natl. Cancer Inst.* **95**, 353–361 (2003).
  66. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
  67. Liao, Y., Smyth, G. K. & Shi, W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* **41**, e108 (2013).
  68. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
  69. Barrett, T. et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* **41**, D991–D995 (2013).
  70. Vizcaino, J. A. et al. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **44**, 11033 (2016).

## Acknowledgements

The authors would like to thank the Genomics core, the staff in the Proteomics Core especially Valar Nila Roamio Franklin and Carmen Gonzalez Tejedo, the Bioinformatics Core especially Dominique Laurent Couturier and Rory Stark, the Histopathology Core especially Jo Arnold and Bev Wilson, the Biorepository Core and the Instrumentation Core facility at Cancer Research UK. We also would like to thank Igor Chernukhin for uploading the RNA-seq data to GEO. The authors would also like to thank Alana Welm for sharing the PDX models and Elena Provenzano for the evaluation of the IHC

staining. We acknowledge the support of the University of Cambridge and Cancer Research UK. We also would like to acknowledge the National Breast Cancer Foundation of Australia (Grant 14001) and the Personalised Medicine for Breast Cancer Biobank, funded by the Sydney Breast Cancer Foundation, National Health and Medical Research Council (NHMRC), TransBCR Centre for research excellence in translational breast cancer research and the McMurtrie family. The Fusion Lumos Orbitrap mass spectrometer was purchased with the support from a Wellcome Trust Multi-user Equipment Grant (Grant # 108467/Z/15/Z). Parts of this work were funded by CRUK core grant (grant numbers C14303/A17197, A19274) to F.M.; Breast Cancer Now Award (grant number 2012NovPR042) to F.M. R.S. is supported by a fellowship from the Novo Nordisk Foundation (NNF 14136). J.S.C. is supported by an ERC Consolidator award and CRUK funding.

## Authors Contributions

E.K.P., R.S., C.S.D.S. and J.S.C., conceptualisation; E.K.P., A.N.H., R.S. and C.S.D.S., methodology; E.K.P., proteomic sample preparation and mass spectrometry; E.K.P., A.N.H. and R.S., RIME experiments; E.K.P., S.O. and R.S., validation assays; E.K.P. and R.S., in vivo experiments; E.K.P., K.K., A.N.H., T.I.R., M.E., R.S. and C.S.R.C., data analysis; K.K. and M.E., development of qPLEXanalyzer; E.K.P., K.K., A.N.H., F.M., M.E., R.S., C.S.D.S. and J.S.C., study design; K.H., K.M.C., A.S. and E.L., collection of human clinical samples and PDX models; R.S., C.S.D.S. and J.S.C., oversaw all experiments; E.K.P., R.S., C.S.D.S. and J.S.C., writing-original draft; All, writing-review, and editing.

## Additional information

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-018-04619-5>.

**Competing interests:** The authors declare no competing interests.

**Reprints and permission** information is available online at <http://www.nature.com/reprints>

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018