

Machine Learning to predict Breast Cancer Biology

Dr Andrew N Holding

Andrew.Holding@York.ac.uk

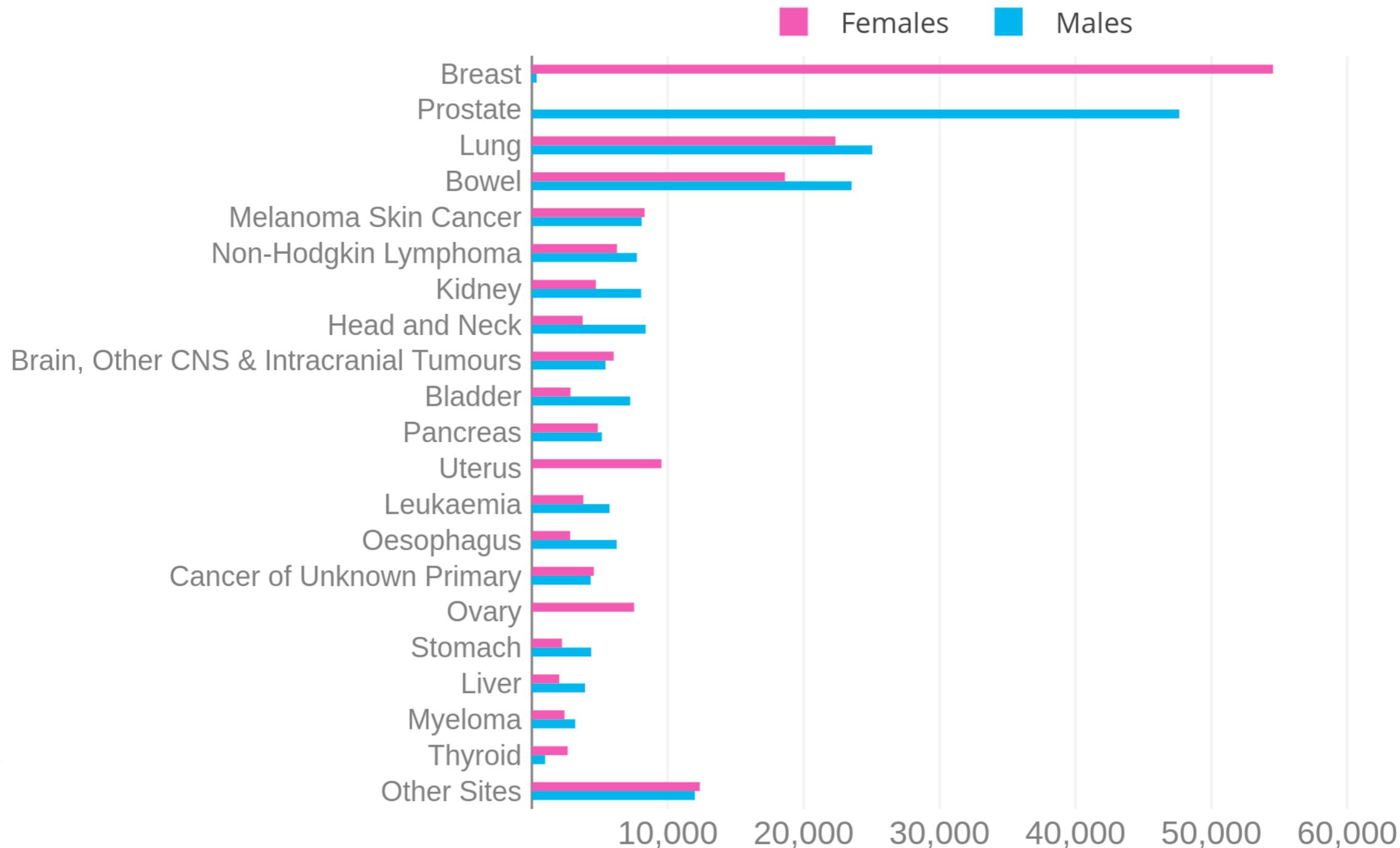
 @AndrewHolding

Turing-Roche Knowledge Share Series, September 2022

Breast Cancer



UNIVERSITY
of York



Source: Cancer Research UK, accessed Jan 2020

<https://www.cancerresearchuk.org/health-professional/cancer-statistics/incidence/common-cancers-compared>

Breast Cancer Subtypes



UNIVERSITY
of York

BREAST CANCER IN WOMEN: KNOW THE SUBTYPE

It's important for guiding treatment and predicting survival.

HR+/HER2- ➤ aka "Luminal A"

73% of all breast cancer cases

- Best prognosis
- Most common subtype for every race, age, and poverty level



HR-/HER2- ➤ aka "Triple Negative"

13% of all breast cancer cases

- Worst prognosis

HR+/HER2+ ➤ aka "Luminal B"

10% of all breast cancer cases

HR-/HER2+ ➤ aka "HER2-enriched"

5% of all breast cancer cases

- Lowest rates for all races and ethnicities

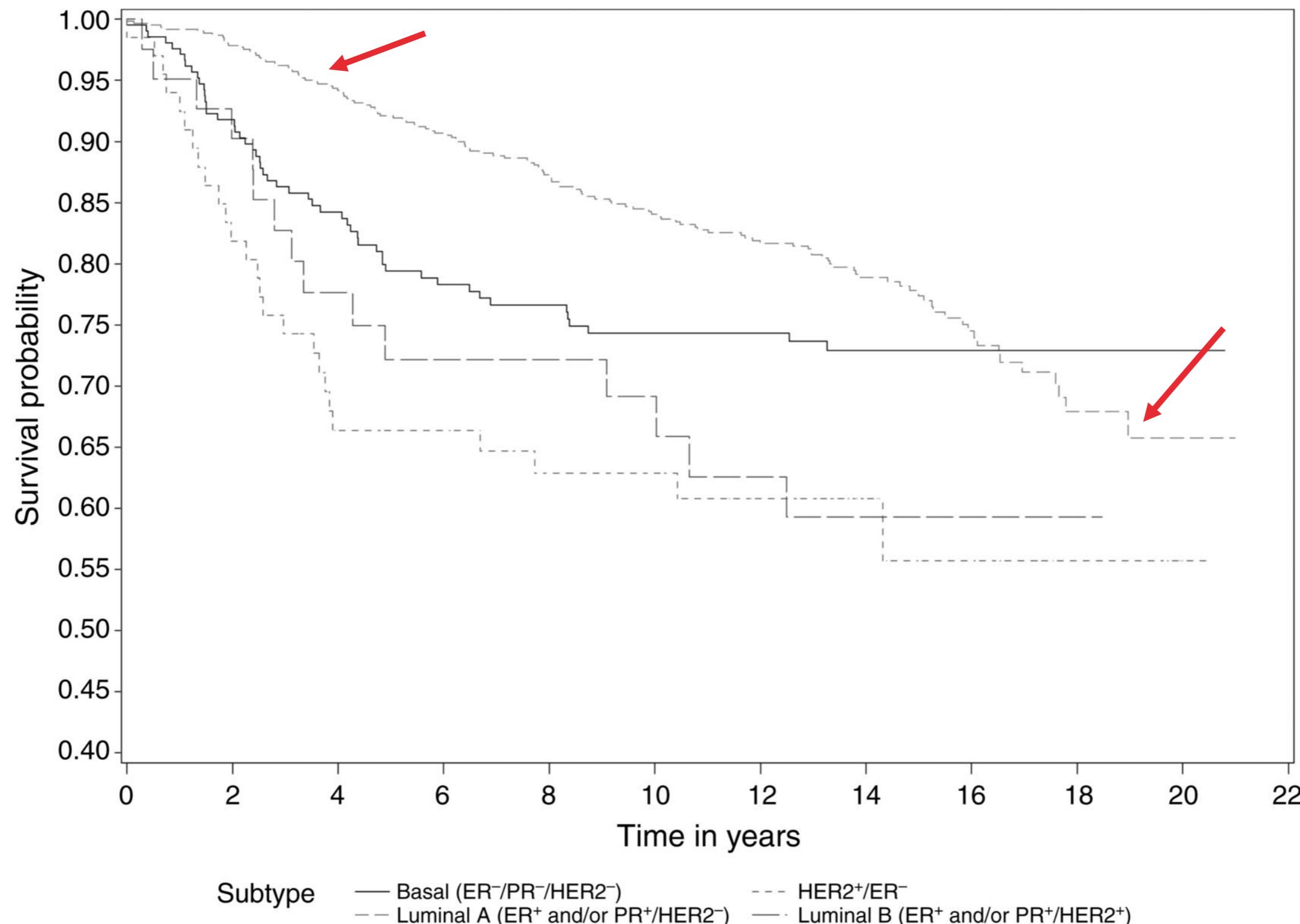
www.cancer.gov

Source: Special section of the Annual Report to the Nation on the Status of Cancer, 1975-2011.

ER+ Breast Cancer Long-Term Relapse Risk is worse than TNBC



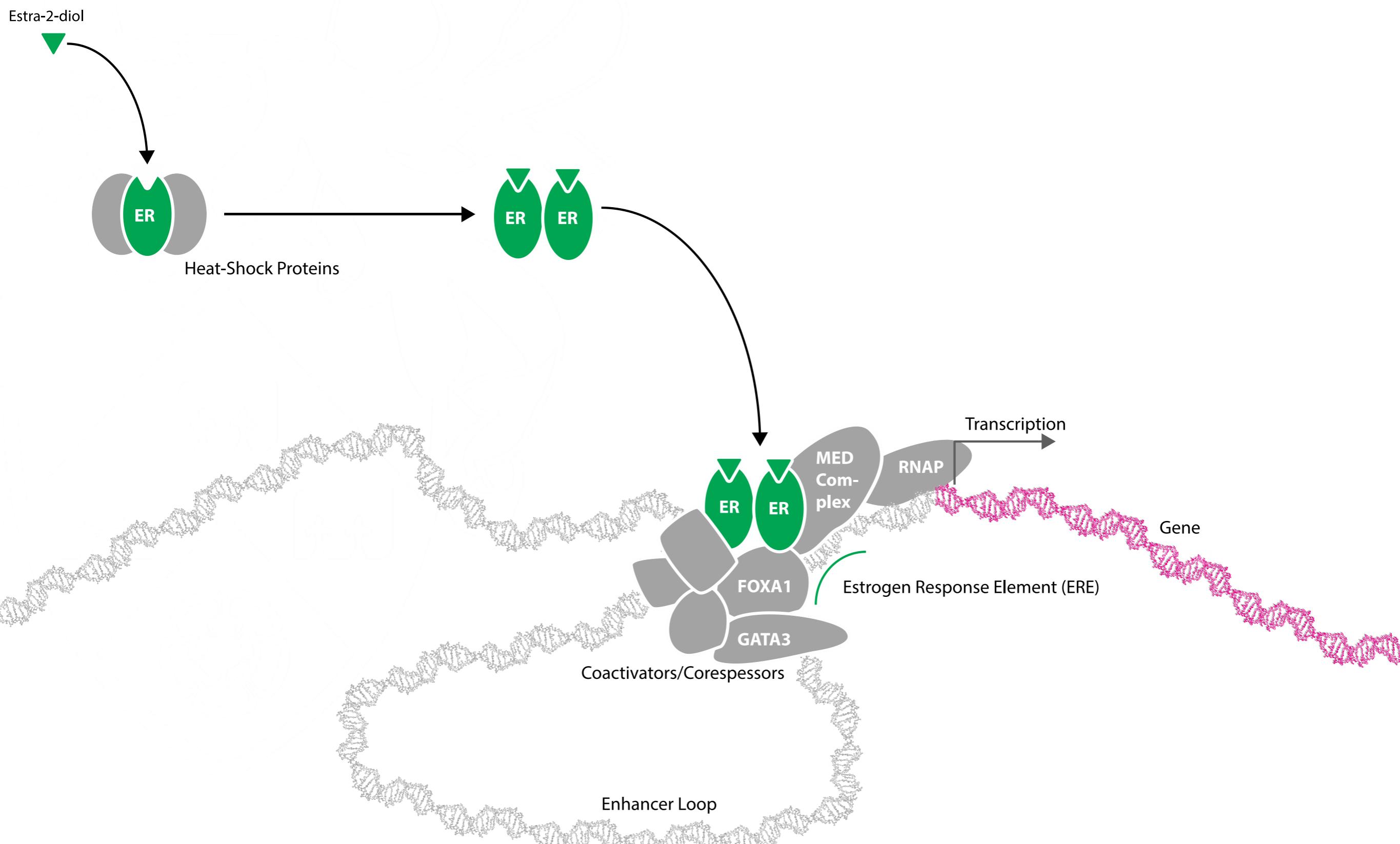
UNIVERSITY
of York



Estrogen Receptor Activates a Transcriptional Programme



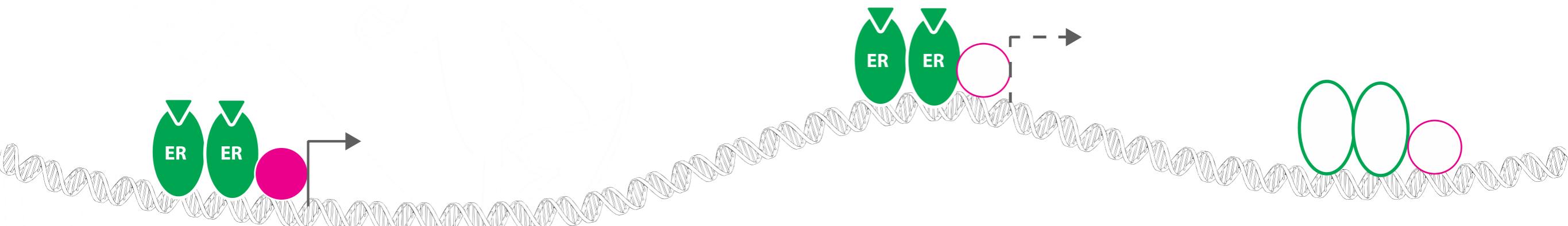
UNIVERSITY
of York



Predicting Protein Activity



Concept: Cofactors that modulate ER activity are themselves at least partly regulated by their own expression.

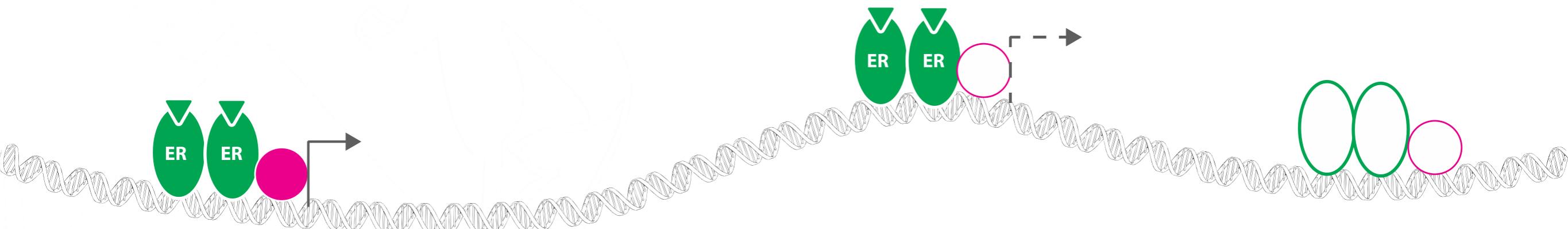


Predicting Protein Activity



UNIVERSITY
of York

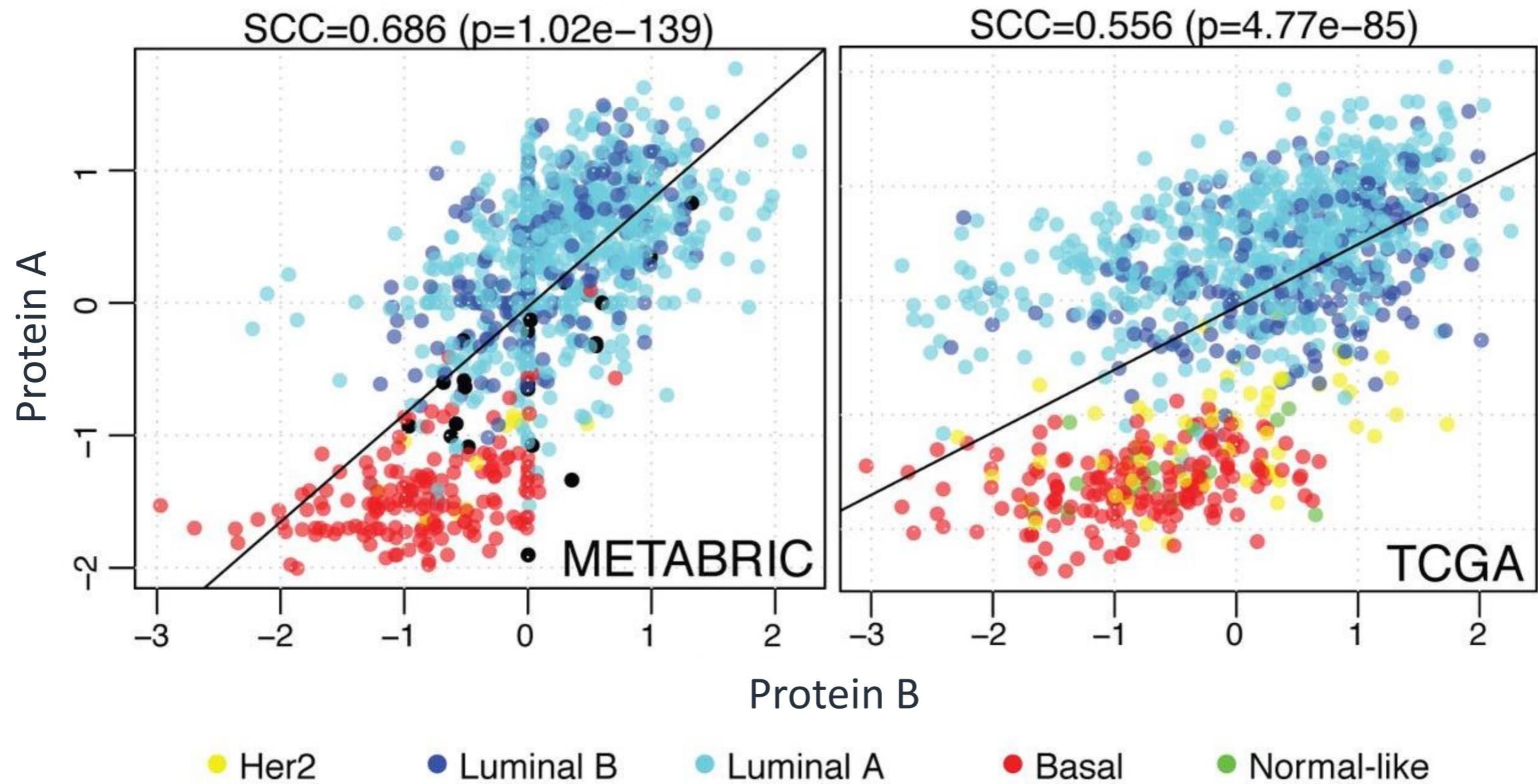
Concept: Cofactors that modulate ER activity are themselves at least partly regulated by their own expression.



Therefore if a cofactor is poorly expressed in a cell the genes it activates will be also less expressed.

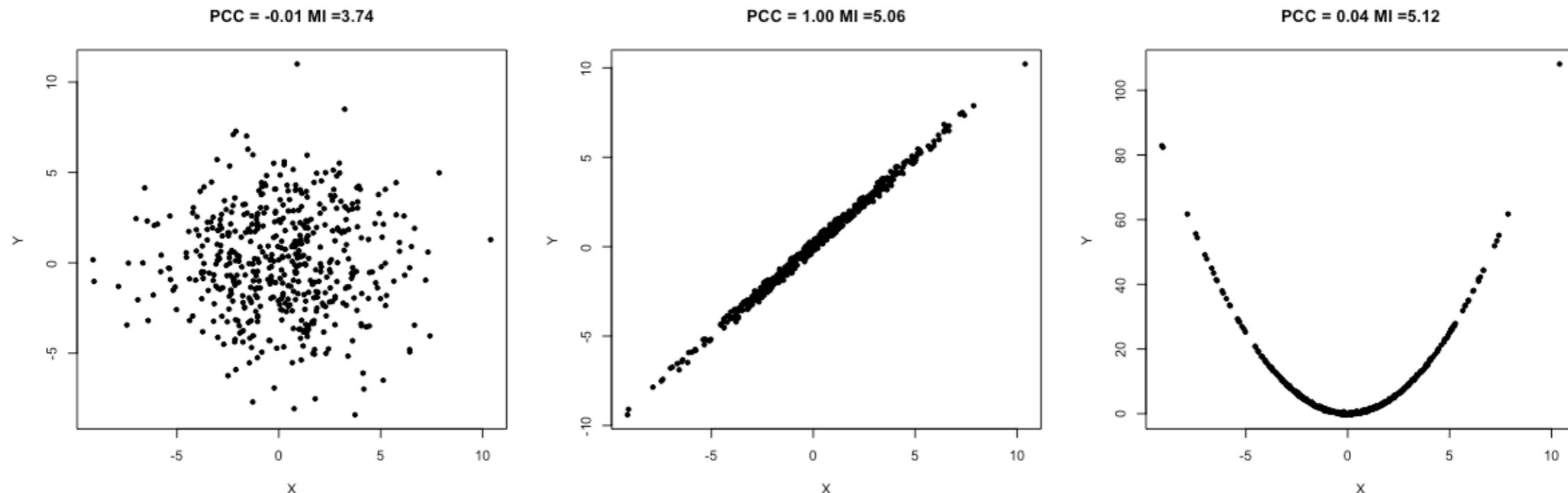


Natural Variance is Plenty

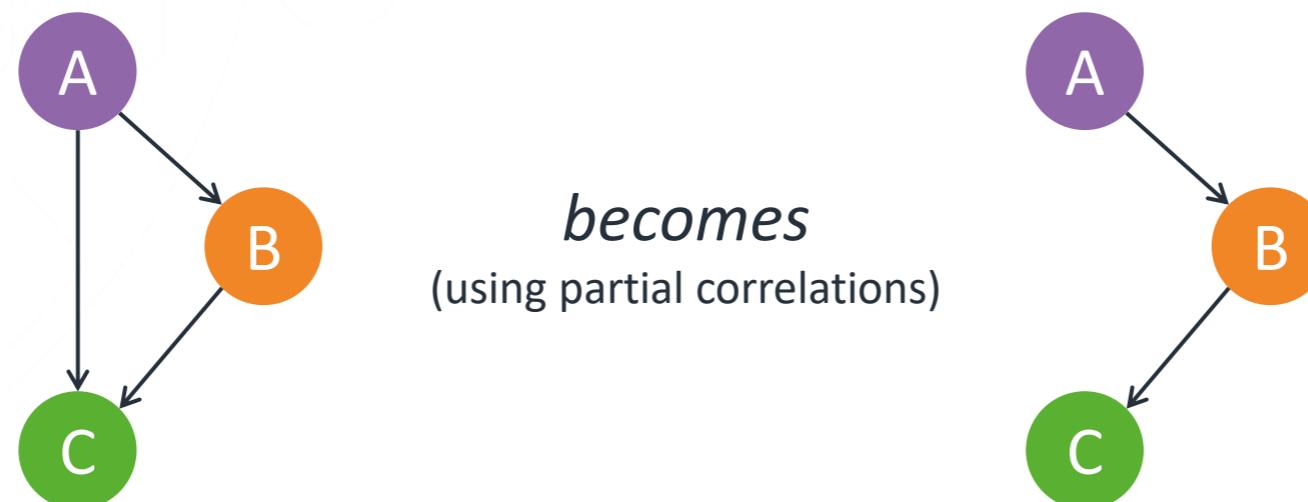


It's a bit more complicated...

- ARACNe-AP uses mutual information over linear regression.



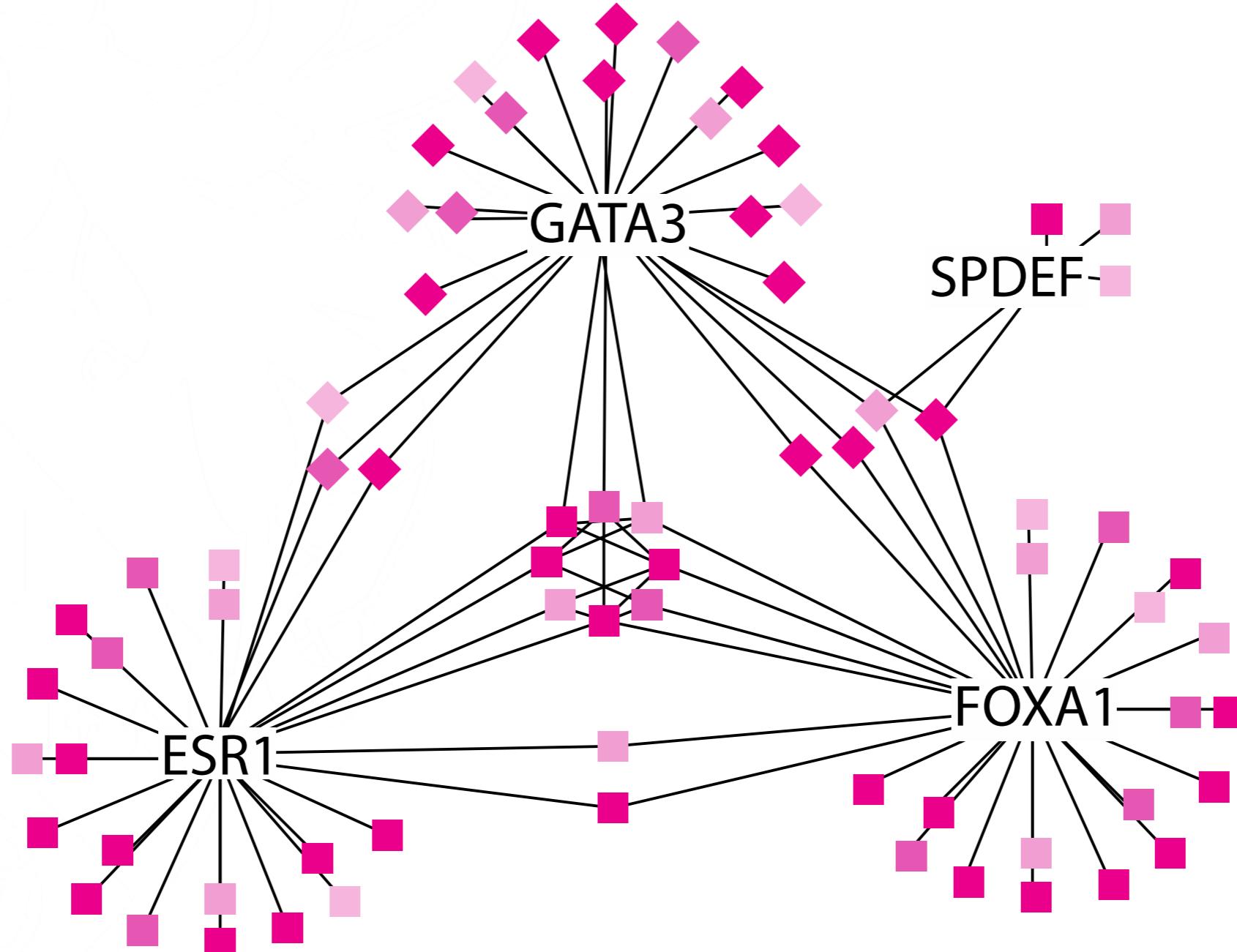
- ARACNe-AP reduces transitive edges.



What is a regulon?



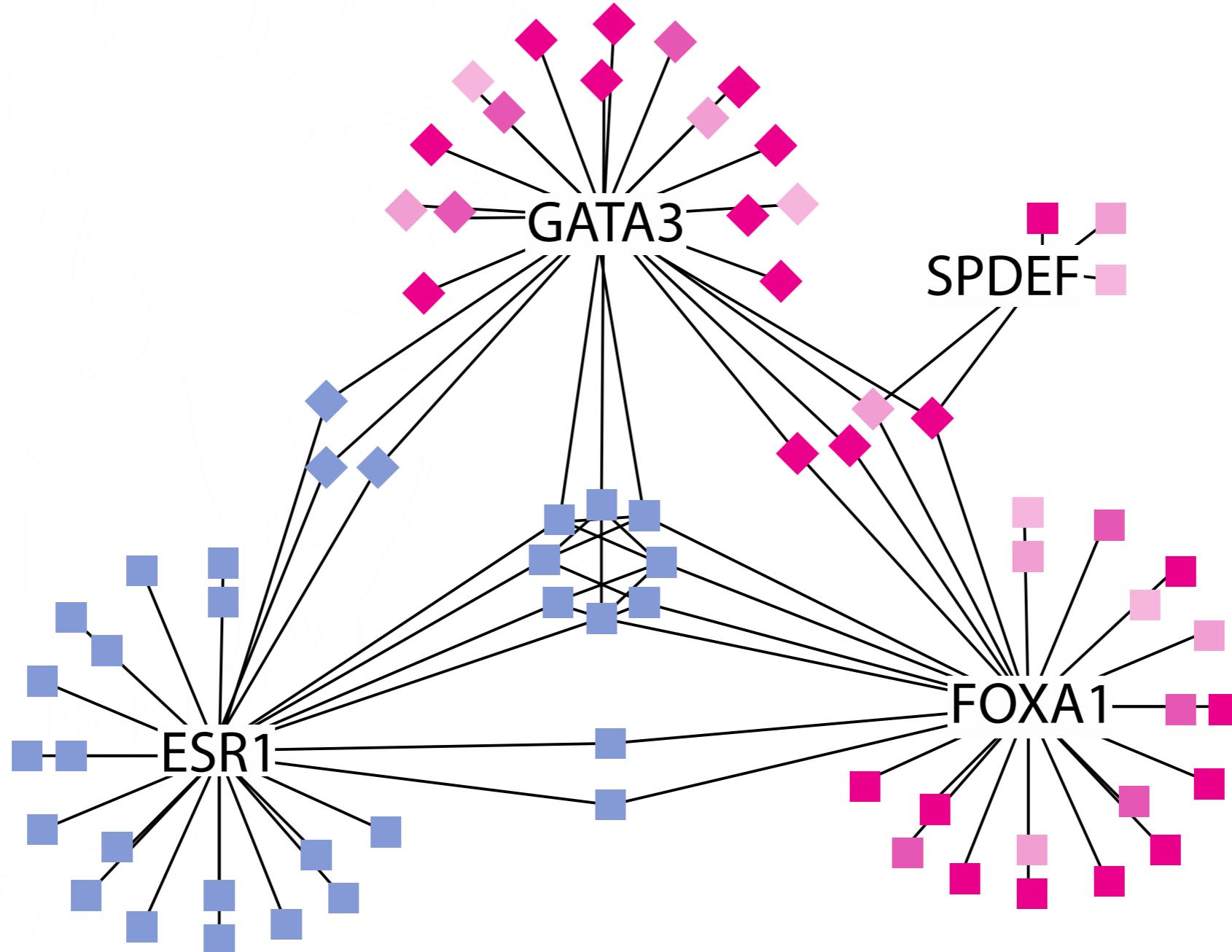
UNIVERSITY
of York



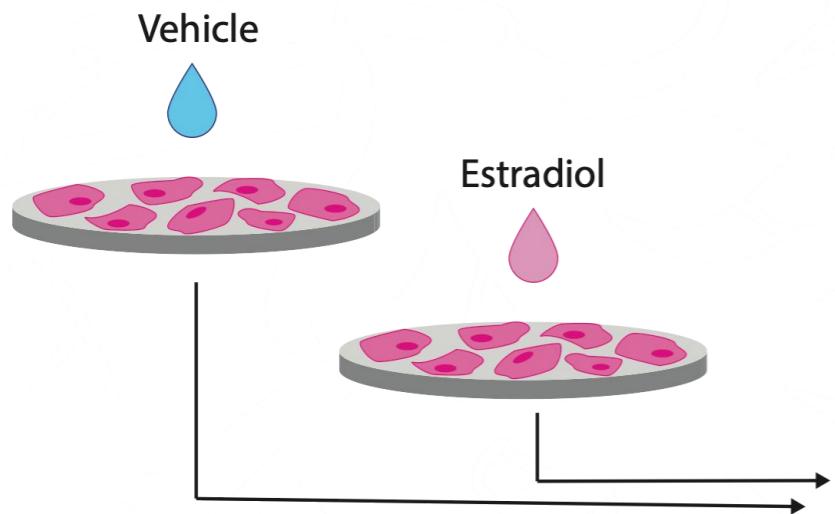
What is a regulon?



UNIVERSITY
of York



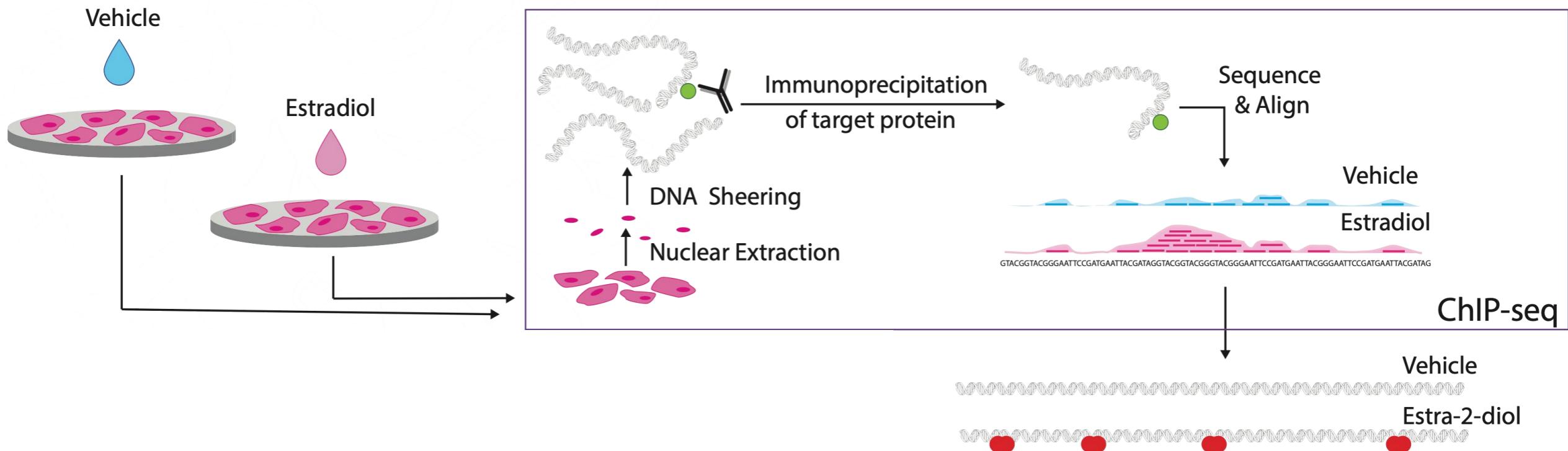
VULCAN



VULCAN



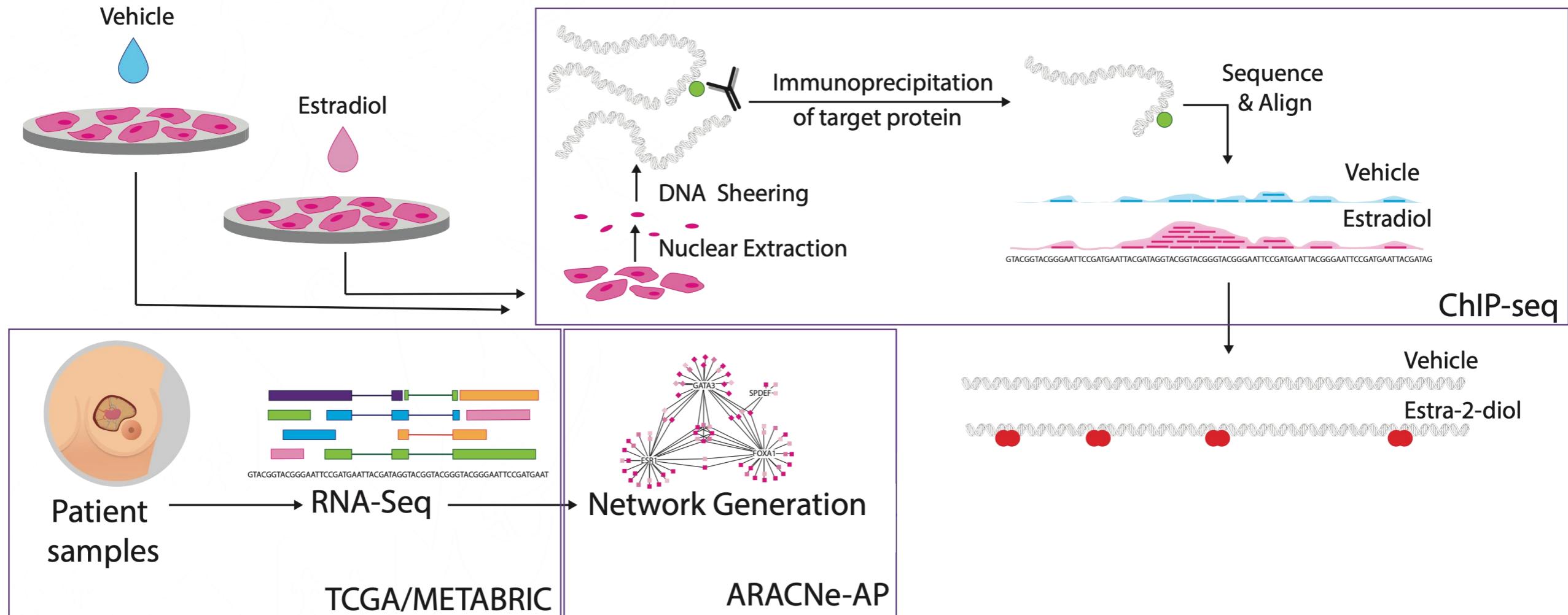
UNIVERSITY
of York



VULCAN



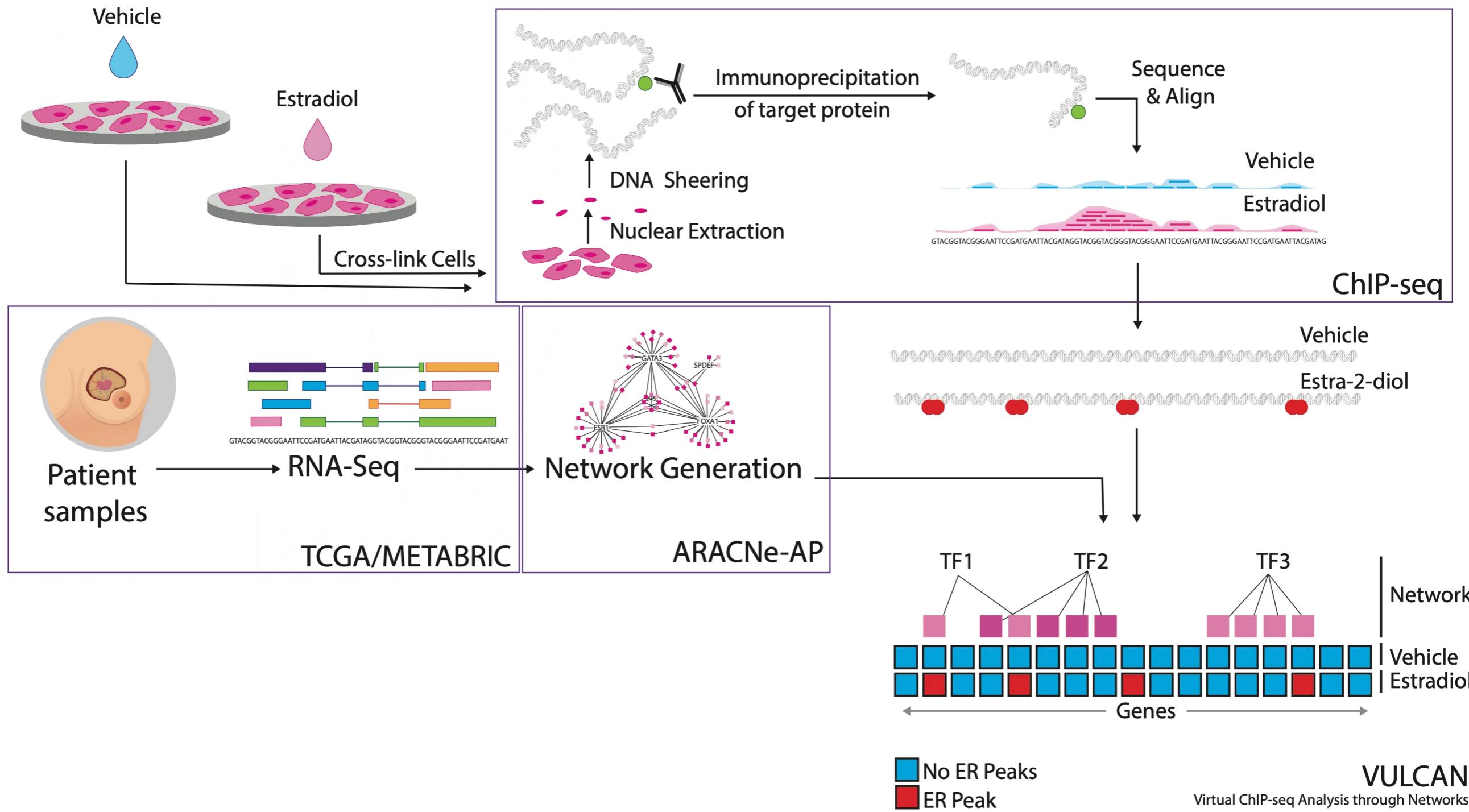
UNIVERSITY
of York



VULCAN



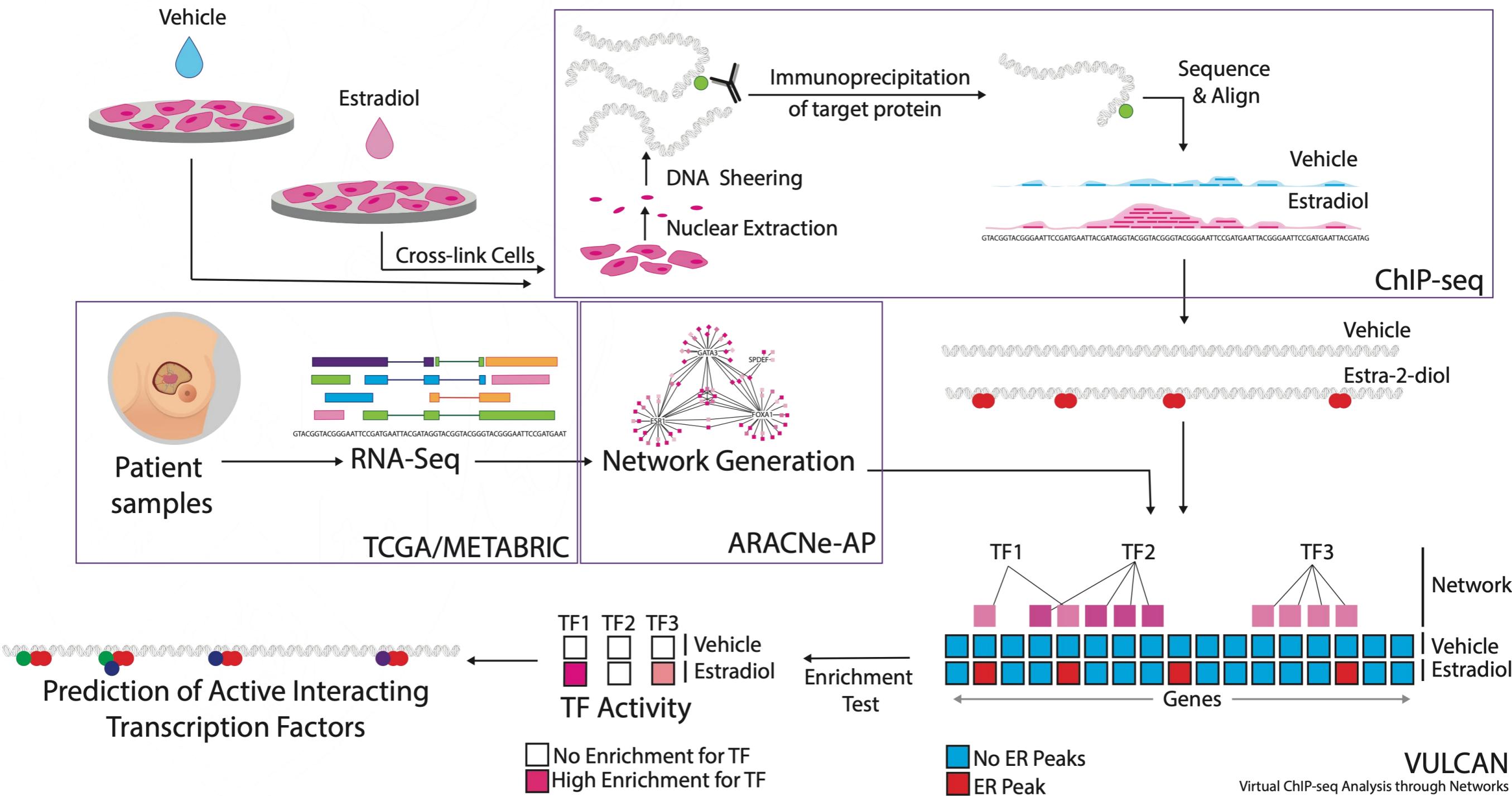
UNIVERSITY
of York



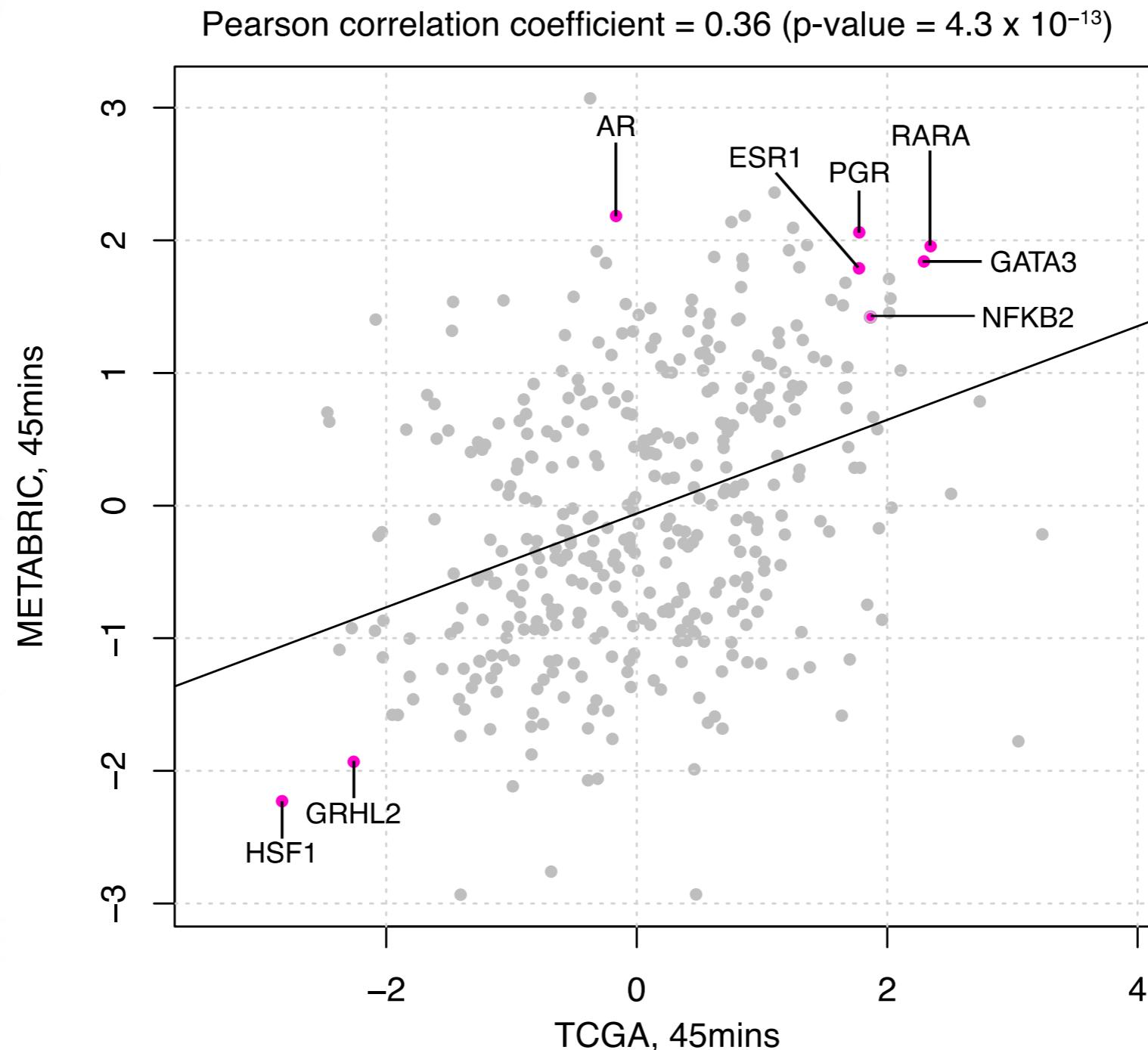
VULCAN



UNIVERSITY
of York



In practice

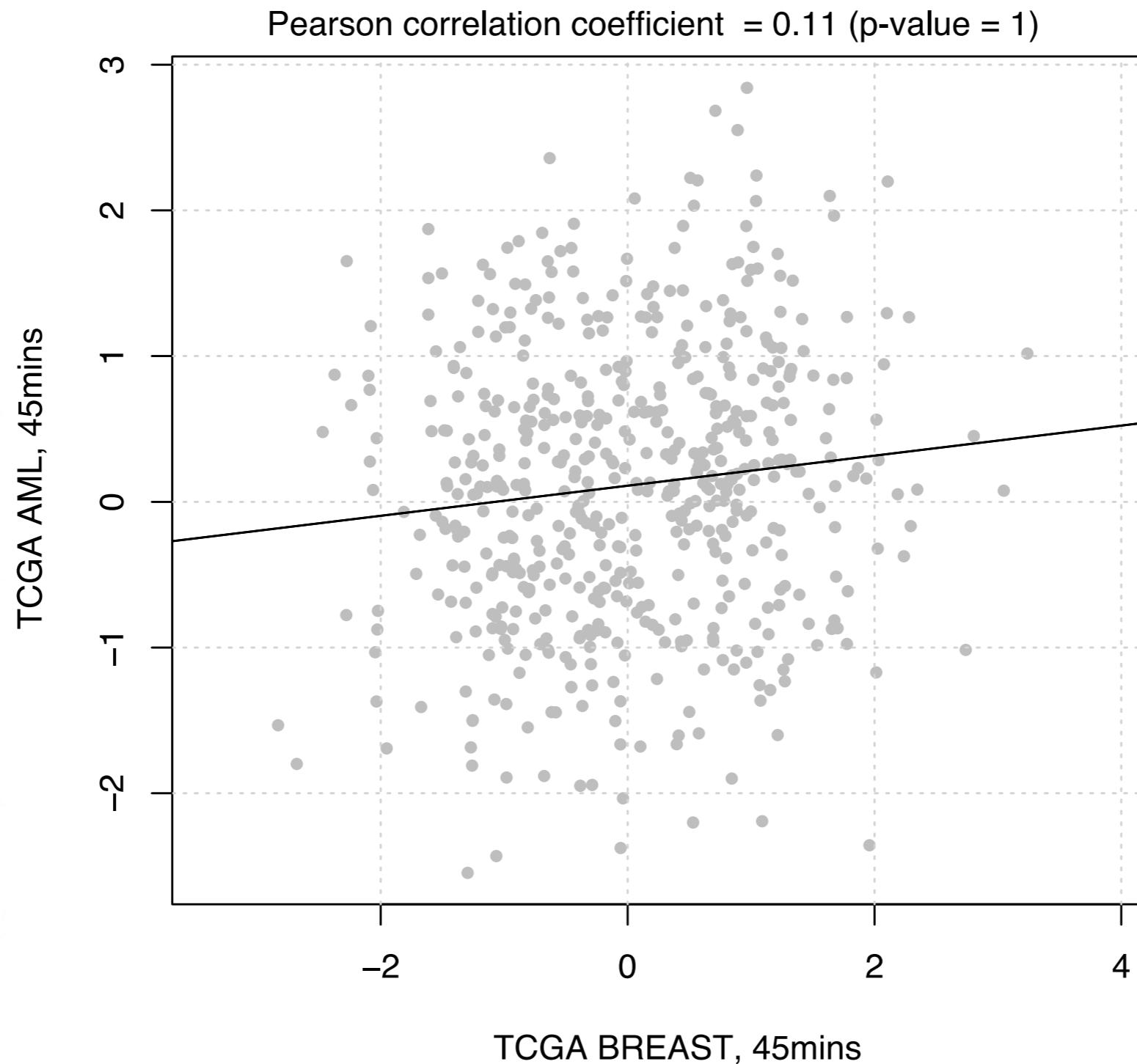


- PGR – Mohammed *et al.*, Nature, 2015
- GATA3 – Carroll *et al.*, Cell, 2005
- RARA – Ross-Innes *et al.* Genes & Dev, 2010
- NFkB2 – Frasor *et al.* Cancer Res., 2009
- AR – Birrell *et al.* J. Steroid Biochem. Mol. Biol. 1995

Negative control

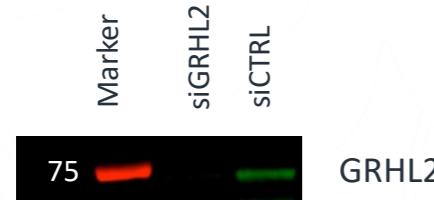


UNIVERSITY
of York

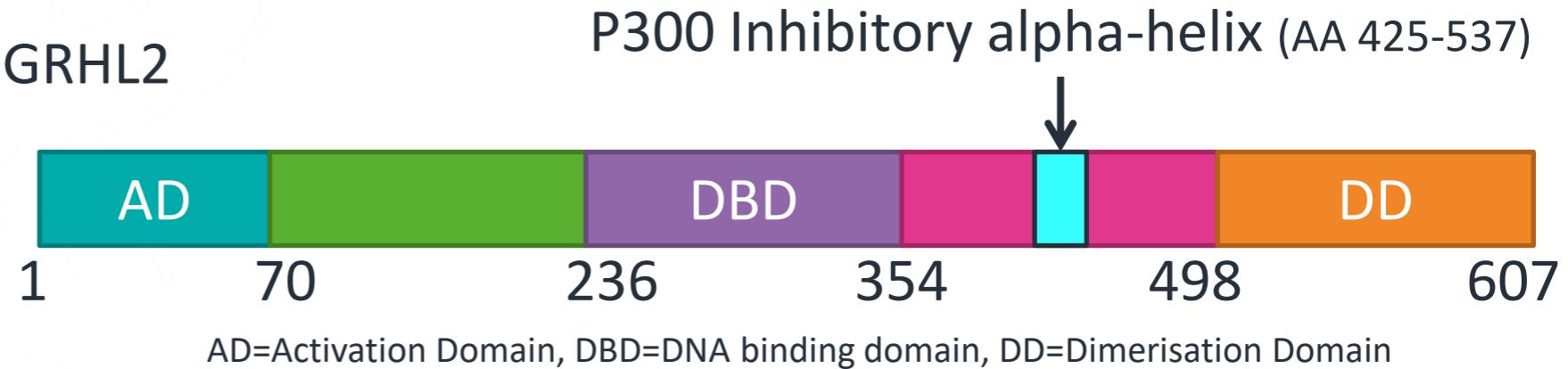


What is GRHL2 doing?

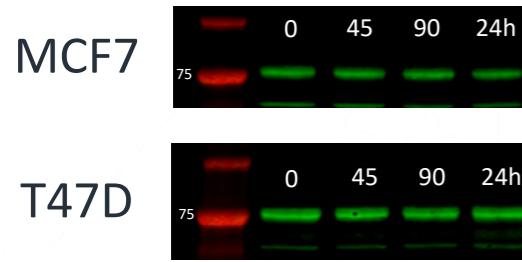
siGRHL2 reduces GRHL2 expression



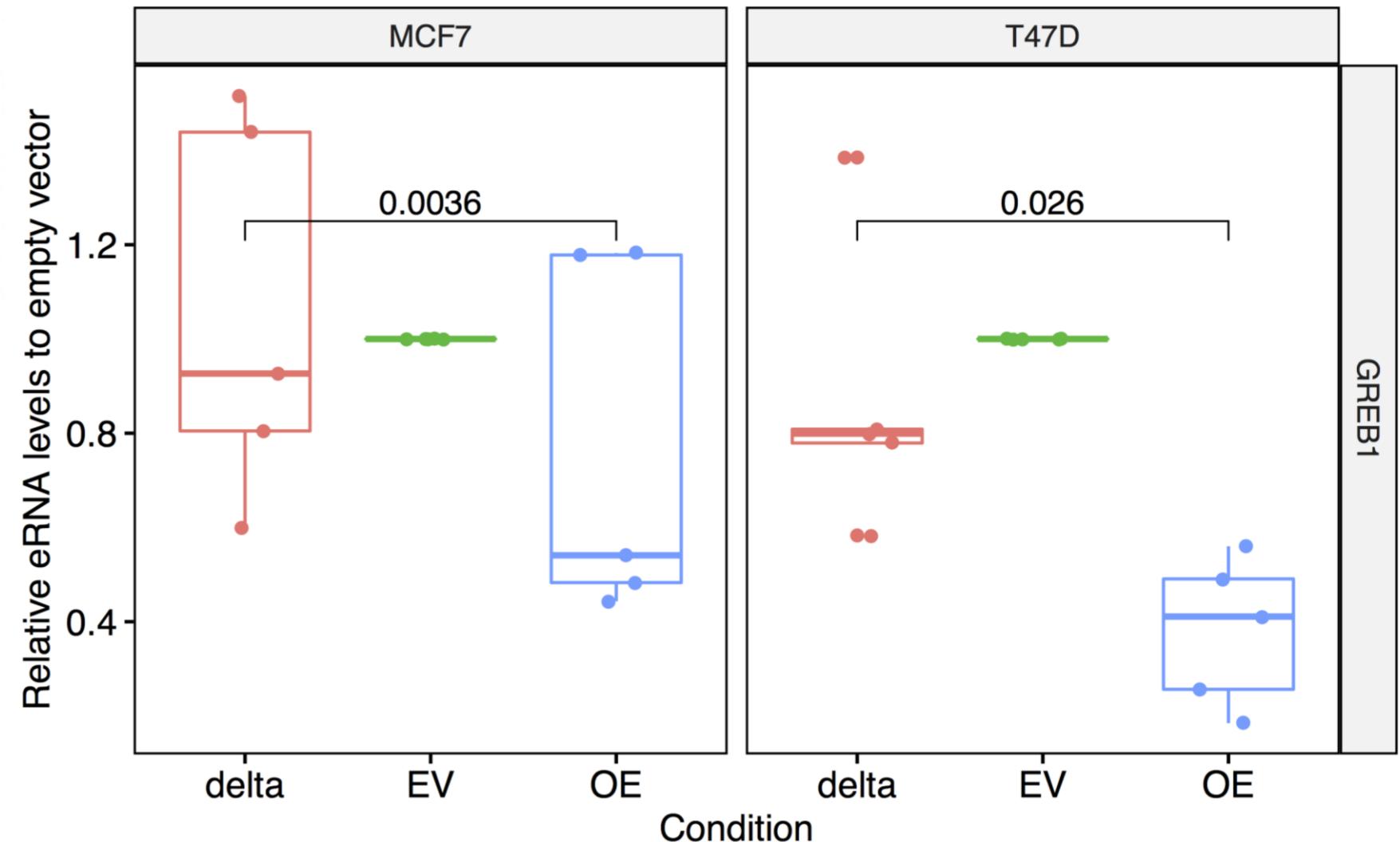
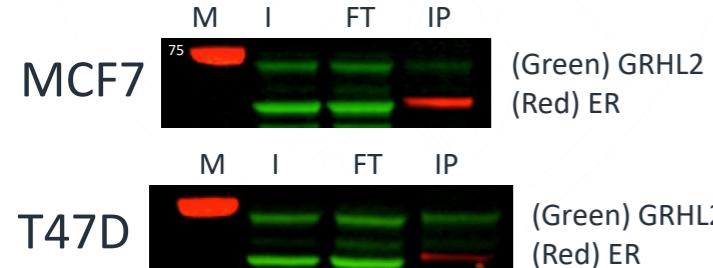
GRHL2



Addition of E2 does not alter GRHL2 expression



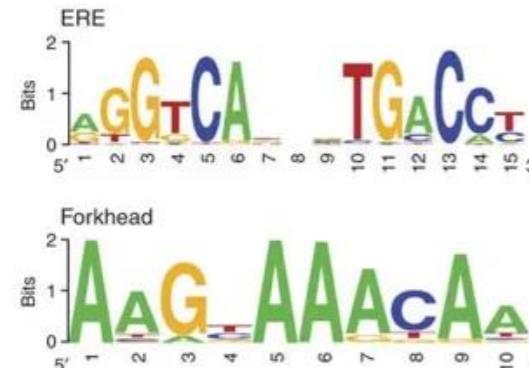
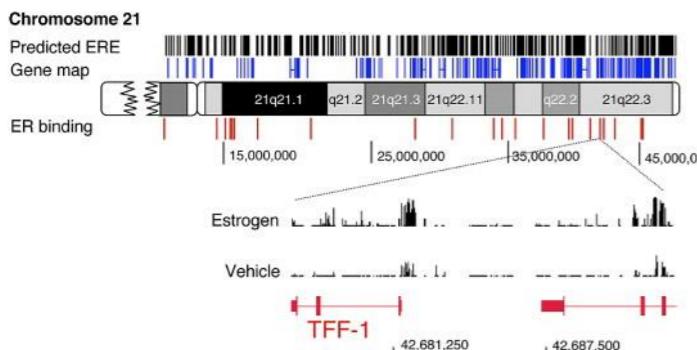
Co-IP of ER and GRHL2 demonstrates protein-protein interactions



Where does VULCAN fit?

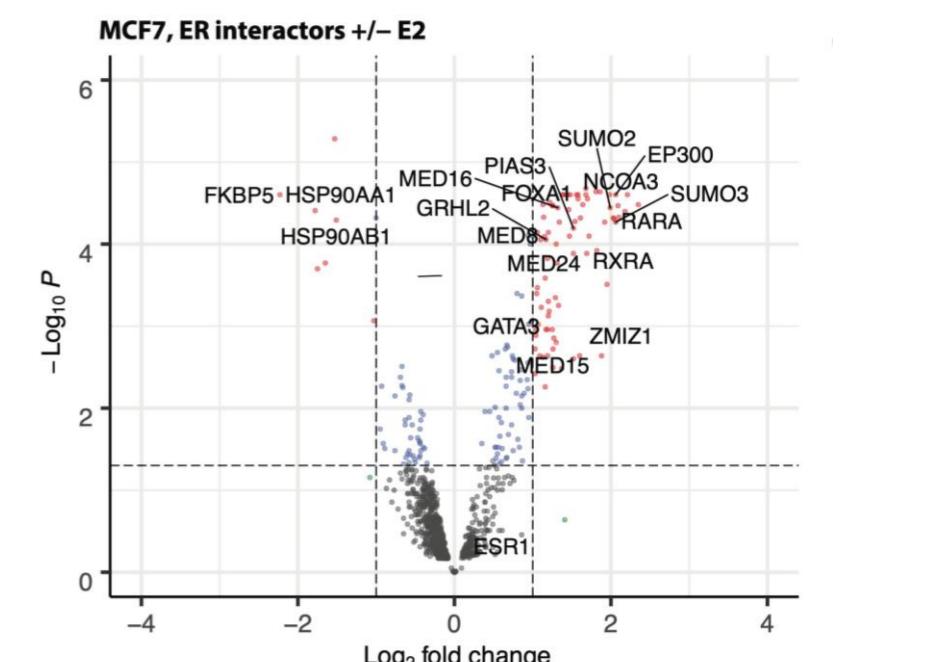
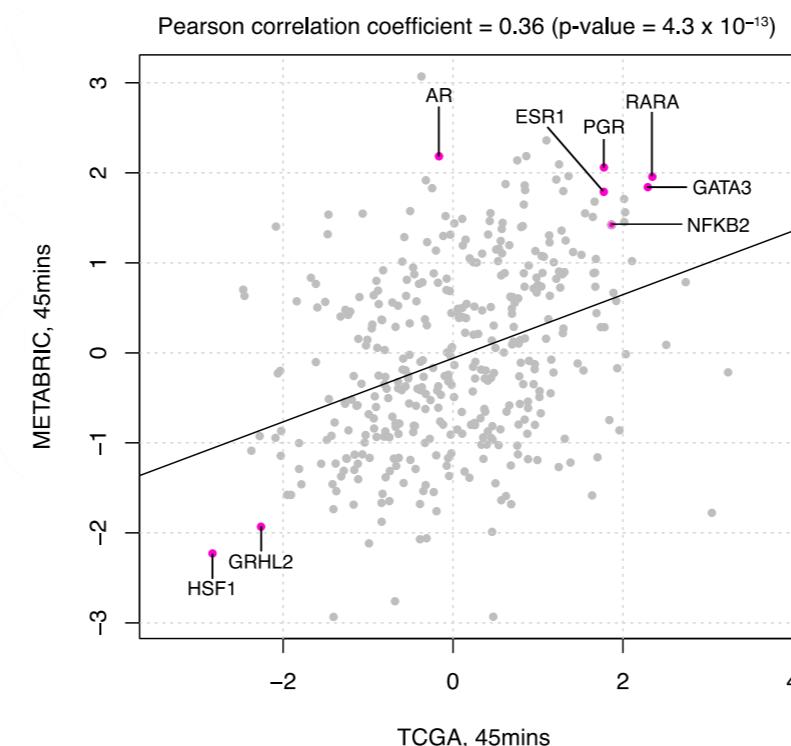


Motif Analysis – What can bind?



Carroll *et al.* Cell 2005

Carroll *et al.* Nature Genetics 2006



Mohammed *et al.* Cell Reports 2013

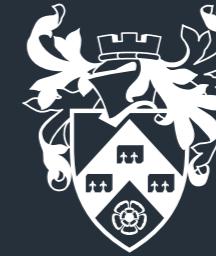
Papachristou *et al.* Nat. Comms. 2018

(Data from Godicelj *et al.* bioRxiv 2019)

VULCAN – What is active (in the tumour)?

Holding *et al.* Genome Biology 2019

Summary



UNIVERSITY
of York

Holding et al. *Genome Biology* (2019) 20:91
<https://doi.org/10.1186/s13059-019-1698-z>

Genome Biology

RESEARCH Open Access

VULCAN integrates ChIP-seq with patient-derived co-expression networks to identify GRHL2 as a key co-regulator of ER α at enhancers in breast cancer

Andrew N. Holding^{1,2†}, Federico M. Giorgi^{1,3†}, Amanda Donnelly¹, Amy E. Cullen¹, Sankari Nagarajan¹, Luke A. Selth⁴ and Florian Markowetz¹

Abstract

Background: VirtuAl ChIP-seq Analysis through Networks (VULCAN) infers regulatory interactions of transcription factors by overlaying networks generated from publicly available tumor expression data onto ChIP-seq data. We apply our method to dissect the regulation of estrogen receptor-alpha activation in breast cancer to identify potential co-regulators of the estrogen receptor's transcriptional response.

Results: VULCAN analysis of estrogen receptor activation in breast cancer highlights the key components of the estrogen receptor complex alongside a novel interaction with GRHL2. We demonstrate that GRHL2 is recruited to a subset of estrogen receptor binding sites and regulates transcriptional output, as evidenced by changes in estrogen receptor-associated eRNA expression and stronger estrogen receptor binding at active enhancers after GRHL2 knockdown.

Conclusions: Our findings provide new insight into the role of GRHL2 in regulating eRNA transcription as part of estrogen receptor signaling. These results demonstrate VULCAN, available from Bioconductor, as a powerful predictive tool.

Keywords: Breast cancer, Network analysis, Dynamics, ER, Master regulator, ChIP-seq, VULCAN, GRHL2, P300, H3K27ac

Introduction
Breast cancer is the most common form of cancer in women in North America and Europe accounting for 31% of all new cancer cases. In the USA, it is estimated that 41,400 deaths will have occurred from the disease in 2018 [1]. The majority of breast cancers, approximately 70%, are associated with deregulated signaling by the estrogen receptor-alpha (ER), which drives tumor growth. Therefore, in ER-positive (ER+) tumors, ER is the primary therapeutic target. During activation, ER recruits several cofactors to form an active complex on the chromatin. FOXA1 is of particular interest as the protein shares nearly 50% of its genomic binding sites with ER and has been shown to operate as a pioneer factor before ER activation [2, 3]. It is through FOXA1 and other cofactors (e.g., SRC-1) [4, 5] that ER is able to recruit RNA polymerase II at the gene promoter sites by way of adaptor proteins in order to initiate transcription [6]. Combinatorial treatments targeting ER cofactors present a significant opportunity for cancer therapy for increasing patient survival. In particular, the pioneer factor FOXA1 has been identified as a novel therapeutic target for the treatment of breast cancer via transcriptional axis regulation [7].

in therapeutic

* Correspondence: anh25@cam.ac.uk.
† Andrew N. Holding and Federico M. Giorgi contributed equally to this work.
1 CRUK Cambridge Institute, University of Cambridge, Robinson Way, Cambridge CB2 0RE, UK
2 The Alan Turing Institute, 96 Euston Road, Kings Cross, London NW1 2DB, UK
Full list of author information is available at the end of the article

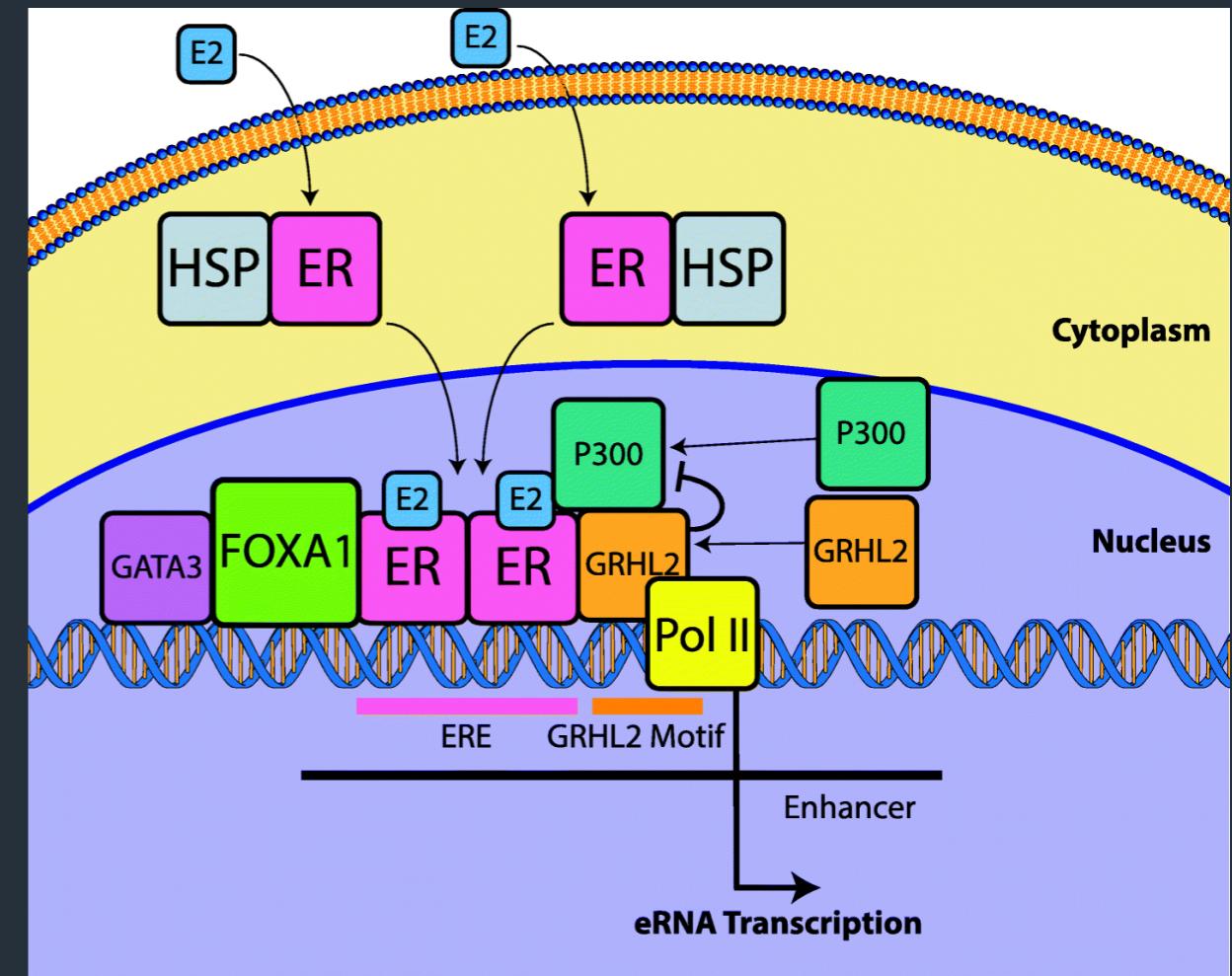
BMC

VULCAN

www.bioconductor.org

VULCAN:

- Predicts interacting TFs at ChIP-seq peaks
- Works on single condition or differential ChIP
- Theoretically works with any TF / Disease
- R-package on Bioconductor
- Published Holding *et al.* *Genome Biology* 2019



Acknowledgements

Holding Lab

Susanne Rose

Wieye Zhao

Jack Stenning

Jane Nesworthy

Heidi Denton

Markowetz Lab

Federico Giorgi

(Now University of Bologna)

Amy Cullen

& the rest of the lab

Carroll Lab

Sankari Nagarajan

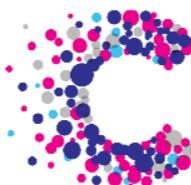
CRUK CI

Luke Selth

Flinders University, Adelaide



Funding



CANCER
RESEARCH
UK



UNIVERSITY OF
CAMBRIDGE

The
Alan Turing
Institute



DOWNING COLLEGE
CAMBRIDGE

breast cancer
now