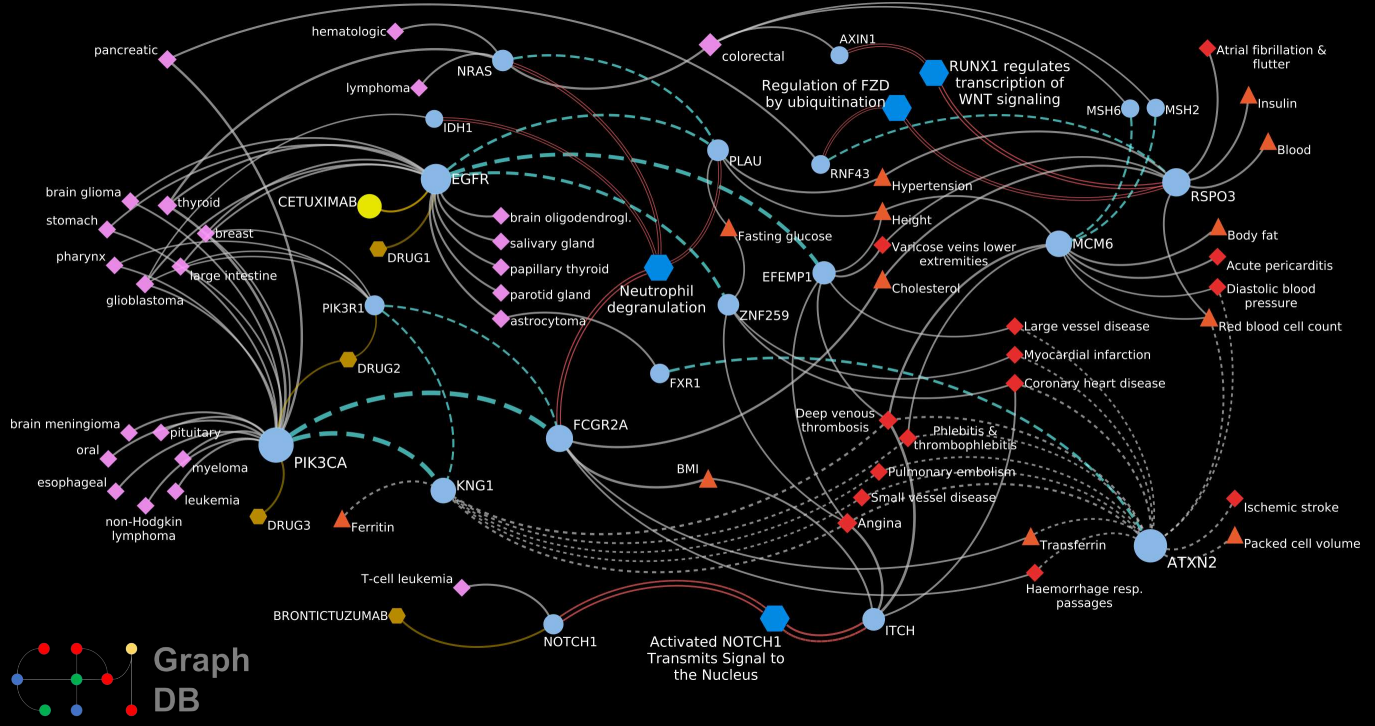

Figure 1. This is a pilot view to explore the potential of EpiGraphDB to inform us about proteins that are linked to the pathophysiology of cancer and cardiovascular disease (CVD). For each cancer type (pink diamonds), we searched for cancer related proteins (light blue circles) that interact with other proteins identified as protein quantitative trait loci (pQTLs) for CVD (red diamonds for pathologies, orange triangles for risk factors). These pQTLs can be acting in cis (solid lines) or trans-acting (dotted lines). Proteins can interact either directly, a protein-protein interaction (dotted blue edges), or through the participation in the same pathway (red parallel lines). Shared pathways are represented with blue hexagons. We also queried which of these proteins are targeted by existing drugs. We found that the cancer drug cetuximab (yellow circle) inhibits EGFR. Other potential drugs are depicted in light brown hexagonal meta-nodes that are detailed below.

Deciphering molecular mechanisms and prioritizing therapeutic targets in cardio-oncology

Pau Erola[1,2], Benjamin Elsworth[1,2], Yi Liu[2], Valeriia Haberland[2] and Tom R Gaunt[1,2,3]

[1] CRUK Integrative Cancer Epidemiology Programme; [2] MRC Integrative Epidemiology Unit, University of Bristol; [3] The Alan Turing Institute

Cancer and cardiovascular disease (CVD) make by far the immense contribution to the totality of human disease burden, and although mortality is declining the number of those living with the disease shows little evidence of change (Bhatnagar *et al.*, Heart, 2016). Last decade the emerging field of cardio-oncology has recognized common risk factors and biological mechanisms that predispose to both pathologies (Meijers and de Boer, Cardiovascular Research, 2019), but very little is known of the downstream causal pathways. The use of the large amount of existing omics data in both cancer and CVD may allow us to understand the molecular networks that contribute to the progression of both malignant tumours and heart disease.

Here we propose to create a framework to reverse engineer gene regulatory networks driving these phenotypes and decipher the effects of genetic and environmental variability. We will use probabilistic graphical models to genotype and transcriptome data of publicly available datasets, and EpiGraphDB to systematically evaluate which proteins and pathways are relevant for both conditions. Data-driven identification and prioritization of key genes will represent a major step in understanding the pathobiology of cancer and CVD, and will guide the development of personalized drugs and therapeutic interventions. In Figure 1 we depicted the results of our exploratory study that aims at identifying potential drug targets and safety issues in current therapies. For it we used the results from phenome-wide association studies of proteins (Zheng et al., bioRxiv 627398, 2019) of diseases and risk factors associated with CVD, and data of relevant drug-gene pairs from the Open Targets Platform.

What is EpiGraphDB?

EpiGraphDB is an analytical platform and graph database that aims to address the necessity of innovative and scalable approaches to harness the increasing complex, high-dimensional epidemiological data and to address research questions of biomedical importance.

The core objectives of the project are to:

- Develop approaches for the appropriate application and interpretation of causal inference in systematic automated analyses of many phenotypes using data from a rich array of bioinformatics resources.
- Apply data mining approaches to the same integrated dataset to make novel discoveries about disease mechanisms and potential interventions relevant to population health

EpiGraphDB is available at **http://www.epigraphdb.org**

Drug meta-nodes. *DRUG1*: nimotuzumab, gefitinib, necitumumab, afatinib, neratinib, erlotinib, lapatinib, panitumumab, osimertinib, brigatinib, vandetanib, rociletinib, olmutinib, depatuxizumab mafodotin, dacomitinib, pyrotinib, zalutumumab, egf816, varlitinib, poziotinib, matuzumab, canertinib, tesevatinib, futuximab, mab-425, imgatuzumab, bms-690514, pelitinib, duligotuzumab, cudc-101, ac-480, cep-32496, theliatinib, aee-788, mp-412, tak-285, mdx-447, epitinib, jnj-26483327, simotinib. *DRUG2*: copanlisib, buparlisib, taselisib, gedatolisib, dactolisib, pf-04691502, pictilisib, sonolisib, apitolisib, sf-1126, rg-7666, ly-3023414, pilaralisib, voxtalisib, omipalisib, vs-5584, panulisib, pa-799, ds-7423, bgt-226, zstk-474, wx-037, gsk-1059615. *DRUG3*: alpelisib, mln-1117, fimepinostat, pki-179, bay-1082439.

Acknowledgments. This project is funded by grants MRC programme MC_UU_00011/4 and CRUK programme C18281/A19169.

MRC Integrative Epidemiology Unit

University of BRISTOL

INTEGRATIVE CANCER EPIDEMIOLOGY PROGRAMME

CANCER RESEARCH UK