MRes Rotation Project 2024

Characterisation of two newly discovered ASXL-related proteins: effects on chromatin structure.

Principal supervisor's name: Prof. **Andrew Firth**
Principal supervisor's email address: aef24@cam.ac.uk

Principal supervisor's CRUK CC theme: Cell and Molecular Biology Programme

Department for student registration: Pathology
Department or institute where research will take place: Addenbrooke's site, Department of Pathology

Co-supervisor's name (if applicable): Dr. Vicente Perez Garcia
Co-supervisor's email address (if applicable): vperez@cipf.es

Postgraduate scheme: **MRes + PhD (1 + 3-year non-clinical applicants only)**

MRes project outline:

The ASXL proteins are transcriptional regulators that play dual roles in both activating and silencing homeotic loci by interacting with chromatin modification complexes. These genes play essential roles during embryogenesis and germline mutations lead to a range of developmental disorders. ASXL1 and ASXL2 also function as tumour suppressors to maintain normal haematopoietic cell functions. Most ASXL1 somatic mutations lead to truncated protein products, with ASXL1 being one of the most commonly disrupted genes in myeloid malignancies.

We recently described an ~150 codon previously unidentified ORF, overlapping the central region of the ASXL coding sequence, present in both ASXL1 and ASXL2 ("TF", TransFrame) (Dinan et al., 2017; Mudge et al., 2019). We predict that the TF polypeptides are expressed as fusions with the N-terminal half of the ASXL proteins, via the mechanism of programmed ribosomal frameshifting (PRF), giving rise to the transframe proteins ASXL1-TF and ASXL2-TF (Dinan et al., 2017). Mass spectrometry data confirms that both TF polypeptides are expressed and also confirms the site of PRF in ASXL1-TF (unpublished data). The truncated ASXL proteins commonly expressed in myeloproliferative neoplasms resemble defective versions of the natural ASXL-TF proteins, whose expression is no longer tightly restricted by PRF.

The MRes project objectives are to:
- Engineer cell lines to inducibly express ASXL1-TF and ASXL2-TF.
- Use these cell lines to characterise the effects of ASXL1-TF and ASXL2-TF expression upon chromatin modifications, using next generation sequencing technologies.

MRes experimental plan:
In the long term, we wish to understand the function of these previously undiscovered human proteins, ASXL1-TF and ASXL2-TF. The starting point is to establish their effects on chromatin structure and the associated gene expression profile. We will use the piggyBac transposon system to make cell lines that inducibly express the proteins of interest (Li et al., 2013). This will allow us to tightly regulate their expression levels. In addition to WT ASXL1-TF and WT ASXL2-TF, we will include mutants of both proteins designed to abrogate predicted protein:protein

interactions. Full length ASXL1 and ASXL2 will also be included as additional controls.

Once the cells are confirmed to express the protein of interest upon induction, we will assess the effects of protein expression upon a range of chromatin modifications. This will allow us to detect differentially accessible chromatin regions. For this, we will utilise Cleavage Under Targets and Tagmentation (CUT&Tag) (Kaya-Okur et al., 2019) and Cleavage Under Targets and Release Using Nuclease (CUT&RUN) (Meers et al., 2019). These recently developed techniques detect and quantify genome-wide specific chromatin modifications. The locations of these modifications can be correlated to gene expression patterns, as RNA-seq may be performed alongside. Notably both techniques require significantly less input sample than CHIP-seq or ATAC-seq, making them more amenable to high-throughput applications. Both have been recently applied to ASXL1 studies by other laboratories (Lin et al., 2023; Braun et al., 2023). A range of chromatin markers will be assessed, including H2AK119Ub, H3K27me3, H3K4me1, H3K4me3, and H3K27Ac.

PhD project outline:

The aims of the PhD project are:
(i) To expand the assessment of the effects of ASXL-TF protein expression on chromatin modifications and gene expression.
(ii) To investigate the PRF expression mechanisms of both ASXL-TF proteins.
(iii) To identify binding partners and elucidate their roles in ASXL-TF function.

Whilst the MRes project aimed to test whether the presence of an ASXL-TF protein correlates with alterations in chromatin modifications, the PhD project will focus upon elucidating the phenotypes caused by natural levels of ASXL-TF expression. Rather than overexpressing the ASXL-TF proteins (such as in the piggyBac system), CRISPR base editing will be utilised to “knock out” endogenous ASXL-TF expression by mutating the PRF site. The effects of ASXL-TF protein depletion upon chromatin structure can then be assessed.

The PhD project will also investigate the expression mechanism of both ASXL-TF proteins. Based on our extensive prior work on PRF-driven gene expression, we predict that the two TF ORFs are accessed via distinct PRF mechanisms (Dinan et al., 2017). In the case of ASXL1-TF, this has been verified by mass spectrometry. However we do not know the PRF efficiencies, or whether RNA or protein stimulators are involved.

Finally, the PhD project will investigate the binding partners of ASXL1-TF and ASXL2-TF. Their shared N terminal domains mean that the ASXL and ASXL-TF proteins are expected to both interact with certain epigenetic regulatory proteins, including BAP1. ASXL1 and BAP1 form a complex that removes H2AK119 marks, relieving transcriptional repression. BAP1 also binds HCF-1 via a non-canonical binding motif. BAP1 may therefore act as a bridging protein, bringing HCF-1 and full-length ASXL proteins into proximity. However, both ASXL1-TF and ASXL2-TF contain an HCF-1 binding motif, not present in the ASXL proteins. This canonical HCF-1 binding motif may compete with the non-canonical site in BAP1, rearranging the trimeric complex and potentially altering its enzymatic activities. These predicted novel interactions mean that the ASXL-TF proteins are likely to be chromatin remodellers with unique functions.

PhD experimental plan:

Importantly, the three Objectives of the PhD project can be performed in parallel, and do not rely upon each other for progression. Selection of a relevant cell line is paramount for the success of downstream experiments and will be conducted in the first year of the PhD. We already have a

selection of candidate cell lines, nominated based on ASXL1/2 transcript levels and a lack of genomic duplications (as many cancer-derived lines that express high levels of ASXL1 do so due to genomic duplications and/or polyploidy, making CRISPR editing extremely difficult). The chosen cell line(s) will be used for both CRISPR editing and elucidating the mechanisms of PRF.

Once the CRISPR cell lines are generated, the effects of the mutations on the chromatin structure and gene expression would then be assessed (years 2 and 3) using NGS techniques (such as CUT&Tag, CUT&RUN, and ChIP-seq). The results would expand upon the preliminary data generated in the MRes project.

To identify protein stimulators, which may drive PRF by binding to nascent mRNA, we will use “RiboTrap” which has been successfully used in the host laboratory to characterise viral PRF (Napthine et al., 2017) (years 2 and 3).

The identification of ASXL-TF protein interaction partners will commence in year 1 and will initially focus on the predicted TF-BAP1-HCF-1 complex. The ASXL-TF proteins will be immunoprecipitated and the bound host partners identified by LC/MS-MS. This will allow us to identify the chromatin remodelling complex(es) to which the ASXL-TF proteins contribute. Confirmation of these partners by multiple approaches (including split luciferase, yeast 2-hybrid and co-immunoprecipitation of tagged proteins) will be conducted in year 2. Targeted mutations to identify interaction sites will be conducted in year 3.

Main techniques:

- CRISPR-based gene editing of multiple cell lines.
- Chromatin analysis: both wet laboratory experiments and next generation sequencing analyses (techniques may include, but are not limited to, RNA-seq, Cut&Run, Cut&Tag, CHIP-seq and ATAC-seq).
- RiboTrap: analysis of RNA-protein complexes.
- Bioinformatics: analysis of high-throughput sequencing datasets.

Key references:

1. Dinan AM, Atkins JF, Firth AE. ASXL gain-of-function truncation mutants: defective and dysregulated forms of a natural ribosomal frameshifting product? Biol Direct. 2017 12:24. PMID: 29037253.

2. Perez-Garcia V, Lea G, Lopez-Jimenez P, Okkenhaug H, Burton GJ, Moffett A, Turco MY, Hemberger M. BAP1/ASXL complex modulation regulates epithelial-mesenchymal transition during trophoblast differentiation and invasion. Elife. 202110:e63254. PMID: 34170818.

3. Lin I, Wei A, Awamleh Z, Singh M, Ning A, Herrera A; REACH Biobank and Registry; Russell BE, Weksberg R, Arboleda VA. Multiomics of Bohring-Opitz syndrome truncating ASXL1 mutations identify canonical and noncanonical Wnt signaling dysregulation. JCI Insight. 2023 8:e167744. PMID: 37053013.

4. Braun TP, Estabrook J, Schonrock Z, Curtiss BM, Darmusey L, Macaraeg J, Enright T, Coblentz C, Callahan R, Yashar W, Taherinasab A, Mohammed H, Coleman DJ, Druker BJ, Demir E, Lusardi TA, Maxson JE. Asxl1 deletion disrupts MYC and RNA polymerase II function in granulocyte progenitors. Leukemia. 2023 37:478-487. PMID: 36526735.

5. Meers MP, Bryson TD, Henikoff JG, Henikoff S. Improved CUT&RUN chromatin profiling tools.

Elife. 2019 8:e46314. PMID: 31232687.

6. Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, Ahmad K, Henikoff S. CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun. 2019 10:1930. PMID: 31036827.

7. Li X, Burnight ER, Cooney AL, Malani N, Brady T, Sander JD, Staber J, Wheelan SJ, Joung JK, McCray PB Jr, Bushman FD, Sinn PL, Craig NL. piggyBac transposase tools for genome engineering. Proc Natl Acad Sci U S A. 2013 110:E2279-87. PMID: 23723351.

8. Napthine S, Ling R, Finch LK, Jones JD, Bell S, Brierley I, Firth AE. Protein-directed ribosomal frameshifting temporally regulates gene expression. Nat Commun. 2017 8:15582. PMID: 28593994.

9. Mudge JM, Jungreis I, Hunt T, Gonzalez JM, Wright JC, Kay M, Davidson C, Fitzgerald S, Seal R, Tweedie S, He L, Waterhouse RM, Li Y, Bruford E, Choudhary JS, Frankish A, Kellis M. Discovery of high-confidence human protein-coding genes and exons by whole-genome PhyloCSF helps elucidate 118 GWAS loci. Genome Res. 2019 29:2073 PMID: 31537640.