AI-Driven Decoding of the Male Regulatory Genome: Predicting Infertility and Ageing Trajectories

Supervisors: Juan M Vaquerizas (MRC LMS), Dennis Wang (Imperial College), and Steven Niederer (Imperial College)

Infertility affects approximately one in six people worldwide, with male factor disease contributing to nearly 50% of cases[1]. A significant portion of these men suffer from **non-obstructive azoospermia (NOA)**, a condition where sperm production fails completely. Despite advancements in whole-genome sequencing (WGS), the majority of NOA cases remain classified as idiopathic. This is largely because our ability to identify mutations outpaces our capacity to interpret them—especially **non-coding or regulatory variants** that do not change a protein but instead disrupt the complex "instruction manual" required for sperm development[2].

Part of the challenge of understanding male fertility lies in the high dimensionality of the genome and the dramatic **3D chromatin remodelling** that occurs during meiosis[3]. Traditional methods struggle to scale these complexities. To address this, we are launching a transdisciplinary project that merges **AI-driven computational modelling** with functional experimental biology to identify the genetic networks controlling male fertility.

The Project

The primary goal of this PhD is to develop and apply an AI framework to predict how genetic variants and ageing impact male reproductive health. This project will integrate two cutting-edge computational approaches:

1. **3D Genomic Modelling:** Using deep learning (e.g., *ChromaFold*), the student will predict 3D chromatin contact maps from single-cell ATAC-seq data[4]. This allows us to map how distant regulatory elements interact with meiotic genes across the spermatogenic trajectory.
2. **AI-Guided Variant Prioritisation:** Working with **Dennis Wang**, the student will develop machine-learning pipelines (including Graph Neural Networks and Bayesian modelling) to assess the impact of rare non-coding variants identified in clinical cohorts[5].

Validated predictions will then feed into a **"Digital Twin" of the male germline**, developed in collaboration with **Steven Niederer**. This digital model will simulate how molecular changes—driven by ageing or metabolic stress—impact fertility outcomes[6]. High-priority candidates will be functionally validated using **CRISPR/Cas9** in *C. elegans* and in vitro human germ-cell models.

The student will join the **LMS Team Science initiative**, a highly collaborative environment involving a team of three postdoctoral researchers and two additional PhD students. This project bridges the LMS's expertise in genome biology (Vaquerizas), AI and genomics (Wang), and biophysical modelling (Niederer). The student will also benefit from direct access to state-of-the-art facilities, world-class expertise in epigenetics and meiosis, and links to clinical collaborators.

We seek a highly motivated student with a background in **computational biology, machine learning, or human genetics** and a strong interest in reproductive health. You will receive world-class training in both AI-driven discovery and experimental validation, equipping you for leadership in the future of precision medicine.

References (max 6) - *optional*:

1. **Agarwal, A.** et al. Male infertility. *Lancet* 397, 319-333 (2021). https://doi.org/10.1016/s0140-6736(20)32667-2

2. **Gaulton, K. J.** et al. Interpreting non-coding disease-associated human variants using single-cell epigenomics. *Nat Rev Genet* 24, 516–534 (2023). https://doi.org/10.1038/s41576-023-00598-6

3. **Kitamura, Y., Namekawa, S. H.** The 3D genome during germline development and meiosis. Trends Genet. (2025). https://doi.org/10.1016/j.tig.2025.11.004

4. **Gao, V. R.** et al. ChromaFold predicts the 3D contact map from single-cell chromatin accessibility. *Nat Commun* 15, 9432 (2024). https://doi.org/10.1038/s41467-024-53628-0

5. **Venkatesh, S. S.** et al. Genome-wide analyses identify 25 infertility loci. *Nat Genet* 57, 1107–1118 (2025). https://doi.org/10.1038/s41588-025-02156-8

6. **Niederer, S. A.** et al. Creation and application of virtual patient cohorts of heart models. *Philos Trans A Math Phys Eng Sci.* 378, 2173 (2020). https://doi.org/10.1098/rsta.2019.0558