A103 · Developmentally Informed AlphaGenome Modeling to Prioritize Noncoding Variants in Genetically Unresolved Congenital Heart Disease

Presenter: Kristine Yang — Harvard Medical School · Boston Children's Hospital

Authors: Kristine T. Yang, Piotr Sliz, Sarah U. Morton

Session: Day 2 · Fri Oct 2, 2026 · 2:15–4:15 PM

Congenital heart disease (CHD) is the most common major birth defect, affecting 1 in 100 newborns. Though most CHD is believed to have a genetic basis, fewer than half of affected individuals receive a genetic diagnosis. Noncoding variants disrupting cardiac regulatory elements during development may contribute to this missing heritability, but systematic interpretation remains challenging. Moreover, most noncoding variant effect prediction models are trained on tissue-level adult datasets, which do not capture the dynamic and cell-type specific nature of cardiogenesis.

To address these gaps, we developed a customized AlphaGenome model trained on single-cell RNA-seq, single-cell ATAC-seq, ChIP-seq, and Hi-C data from human developmental datasets, generating chromatin accessibility and expression tracks resolved by cardiac lineage (atrial/ventricular cardiomyocytes, endocardial/endothelial cells, fibroblasts, and mural cells) and developmental stage (42-152 days post-conception; 0-30 days iPSC-CM differentiation). On held-out genomic regions, predicted accessibility recovered the observed signal with Pearson r = 0.63-0.90, scaling with per-track pseudobulk sample size. We then tested variant effect prediction against 2,463 fine-mapped pediatric cardiac eQTLs (SuSiE PIP ≥ 0.5). All 48 expression tracks predicted measured allelic fold change above chance (Spearman ρ = 0.075–0.204, all FDR < 0.05; directional accuracy 52.4–56.9%, binomial $p < 1\times10^{-5}$), demonstrating the model recovers genuine regulatory variant effects. By establishing a framework for noncoding variant interpretation in CHD, we are now scoring de novo and rare noncoding variants from patients with CHD, integrating predictions with a multi-omic fetal cardiac enhancer and expression atlas to prioritize candidate functional variants.

Developmentally Informed AlphaGenome Modeling to Prioritize Noncoding Variants in Genetically Unresolved Congenital Heart Disease

Kristine T. Yang*[1,2], Piotr Sliz[1,2], Sarah U. Morton[1,2]
[1] Boston Children's Hospital, Boston, MA, USA, [2] Harvard Medical School, Boston, MA, USA

Congenital heart disease (CHD) is the most common major birth defect, affecting 1 in 100 newborns. Though most CHD is believed to have a genetic basis, fewer than half of affected individuals receive a genetic diagnosis. Noncoding variants disrupting cardiac regulatory elements during development may contribute to this missing heritability, but systematic interpretation remains challenging. Moreover, most noncoding variant effect prediction models are trained on tissue-level adult datasets, which do not capture the dynamic and cell-type specific nature of cardiogenesis.

To address these gaps, we developed a customized AlphaGenome model trained on single-cell RNA-seq, single-cell ATAC-seq, ChIP-seq, and Hi-C data from human developmental datasets, generating chromatin accessibility and expression tracks resolved by cardiac lineage (atrial/ventricular cardiomyocytes, endocardial/endothelial cells, fibroblasts, and mural cells) and developmental stage (42-152 days post-conception; 0-30 days iPSC-CM differentiation). On held-out genomic regions, predicted accessibility recovered the observed signal with Pearson r = 0.63-0.90, scaling with per-track pseudobulk sample size. We then tested variant effect prediction against 2,463 fine-mapped pediatric cardiac eQTLs (SuSiE PIP $\geq$ 0.5). All 48 expression tracks predicted measured allelic fold change above chance (Spearman ρ = 0.075–0.204, all FDR < 0.05; directional accuracy 52.4–56.9%, binomial $p < 1\times10^{-5}$), demonstrating the model recovers genuine regulatory variant effects. By establishing a framework for noncoding variant interpretation in CHD, we are now scoring *de novo* and rare noncoding variants from patients with CHD, integrating predictions with a multi-omic fetal cardiac enhancer and expression atlas to prioritize candidate functional variants.