

Research article

HeartMAP: A multi-chamber spatial framework for cardiac cell-cell communication

Tumo Kgabeng^a, Lulu Wang^{a,b}, Harry Ngwangwa^a, Thanyani Pandelani^{a,*}^a Unisa Biomedical Engineering Research Group, Department of Mechanical, Bioresources, and Biomedical Engineering, School of Engineering, College of Science, Engineering, and Technology, University of South Africa, Florida Science Campus, South Africa^b Department of Engineering, Reykjavik University, Reykjavik, Iceland

ARTICLE INFO

Keywords:

Single-cell RNA-seq
Cell-cell communication
Cardiac chambers
Spatial transcriptomics
Therapeutic targets

ABSTRACT

Understanding cell-cell communication within and between the four distinct cardiac chambers is fundamental to elucidating cardiac function and disease mechanisms. Each chamber exhibits unique cellular and molecular characteristics that reflect specialised physiological roles, yet existing frameworks for mapping chamber-specific intercellular networks have remained limited. Here, we present HeartMAP (Heart Multi-chamber Analysis Platform), a computational framework that infers cardiac cell-cell communication networks at chamber resolution through integration of single-cell RNA-seq co-expression patterns and ligand-receptor interaction databases. Using a dataset of 287,269 cells from seven healthy human heart donors (Single Cell Portal SCP498), we identified chamber-specific cell populations, communication networks, and therapeutic targets. HeartMAP employs a progressive three-tier analytical approach comprising basic pipeline analysis, advanced communication modelling and multi-chamber atlas construction to reveal both conserved and chamber-specific signalling pathways; cross-chamber correlation analysis demonstrated the highest similarity between ventricles ($r = 0.985$) and the lowest between left atrium and left ventricle ($r = 0.870$), reflecting functional specialisation. Communication hub analysis identified atrial cardiomyocytes and adipocytes as key signalling centres with hub scores of 0.037–0.047, while differential expression analysis revealed over 150 significantly different genes per chamber pair. These findings establish a molecular foundation for precision cardiology approaches, enabling chamber-specific therapeutic strategies that could improve treatment outcomes for cardiovascular diseases. HeartMAP is freely available as a Python package than can be installed using “pip install heartmap”, the package’s documentation can be found on <https://pypi.org/project/heartmap/>, it can also be accessed via a user-friendly web interface freely available at <https://huggingface.co/spaces/Tumo505/heartmap-cell-analysis>.

1. Introduction

Through specific cellular architectures and molecular programs, the four functionally distinct chambers that make up the human heart work together to maintain circulation [1]. By providing previously unheard-of resolution of cellular heterogeneity and intercellular communication networks, single-cell RNA sequencing has completely transformed cardiac biology [1–3]. Recent developments in computational frameworks such as CellChat have made it possible to conduct systematic analysis of ligand-receptor interactions and signalling pathway dynamics [4,5], while early research established the basic principles of cardiac cell-cell communication through gap junctions, paracrine signalling, and direct

cell contact [6,7]. Chamber-specific gene expression patterns have been recognised since early microarray studies, which identified differential expression of structural proteins, ion channels, and transcription factors between atrial and ventricular tissues [8,9]. These pioneering investigations demonstrated that ventricular chambers exhibit enrichment for MYH7 and metabolic genes, whereas atrial chambers preferentially express genes such as MYL7 and NPPA (atrial natriuretic peptide) [10]. However, technical limitations in isolating and analysing cardiac cells at single-cell resolution continue to restrict systematic analysis of chamber-specific cell-cell communication networks [2,11]. Recent technical advancements have overcome major challenges associated with analysing single cardiac cells. Novel techniques, such as

* Corresponding author.

E-mail addresses: 28346416@mylife.unisa.ac.za (T. Kgabeng), luluw@ru.is (L. Wang), ngwanhm@unisa.ac.za (H. Ngwangwa), epandet@unisa.ac.za (T. Pandelani).<https://doi.org/10.1016/j.csbj.2025.11.015>

Received 9 September 2025; Received in revised form 6 November 2025; Accepted 6 November 2025

Available online 8 November 2025

2001-0370/© 2025 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

single-nucleus RNA sequencing, made it easier to isolate large adult cardiomyocytes, and novel computational tools, such as CellChat, CellPhoneDB, and NicheNet, made it possible to understand complex intercellular signalling networks from expression data [2,5]. These platforms showed that they could find communication hubs, temporal dynamics, and pathway-specific interactions, which changed how we think about cardiac cellular crosstalk [1,5,12].

Frameworks for examining chamber-specific communication patterns across all four cardiac chambers were still lacking despite these advancements. Prior research, such as Zhao et al. (2002), mainly focused on single-chamber analyses or general atrial-ventricular differences, which left important gaps in our knowledge of cross-chamber signalling networks and chamber-specific therapeutic targets [9,13]. Precision medicine approaches showed promise in clinical applications of chamber-specific analysis, especially in conditions like atrial fibrillation, where chamber-specific drug targets could increase therapeutic efficacy [14–16].

In this study, we present HeartMAP (Heart Multi-chamber Analysis Platform), a computationally inferred spatial framework that addresses these limitations through modelling of intercellular communication probabilities and chamber-specific expression gradients, without requiring direct spatial transcriptomic input. Using 287,269 single-cell transcriptomes from healthy human hearts, HeartMAP integrates basic cell biology with advanced communication analysis and chamber-specific characterisation to reveal both conserved and unique signalling networks. This multi-tiered approach enables identification of chamber-specific markers, communication hubs, and therapeutic targets while providing accessible computational tools optimised for standard hardware platforms.

While HeartMAP integrates chamber-level spatial context, it is important to distinguish this framework from true spatial transcriptomics technologies. HeartMAP infers intercellular communication patterns from co-expression relationships and curated ligand-receptor databases within chamber-specific cell populations, rather than measuring spatial distance and physical proximity between cells. The term ‘spatial’ in HeartMAP refers to the chamber-level anatomical organisation, not sub-cellular or intra-tissue spatial transcriptomics resolution. This distinction is critical for accurate interpretation of the framework’s capabilities and appropriate application to future studies employing true spatial transcriptomics platforms.

2. Methods

2.1. Data acquisition and preprocessing

HeartMAP analyses were performed on the “Transcriptional and Cellular Diversity of the Human Heart” dataset (study ID SCP498) hosted on the Broad Institute’s Single Cell Portal, an open-access repository that follows unrestricted download of processed matrices and metadata [17], [18]; raw sequencing files are deposited in dbGaP (the database of Genotypes and Phenotypes) under accession phs001539.v1.p1, permitting full reproducibility [17]. The original study, published by Tucker et al. in *Circulation* (2020) [17], profiled 287,269 single cardiac nuclei from seven non-failing adult donors comprising of four males and three females, obtained with Institutional Review Board approval and informed consent. Tissue from all four chambers was sampled for each donor, yielding balanced representation: right atrium (RA) 28.4 %, left ventricle (LV) 27.0 %, left atrium (LA) 26.4 %, and right ventricle (RV) 18.2 %; nuclei were isolated to circumvent adult cardiomyocyte size constraints, bar-coded with the 10x Genomics 3’ v3 platform, and sequenced to a mean depth of 50,000 reads per nucleus [17].

Raw count matrices were imported into the HeartMAP pipeline and subjected to strict filtering:

- Mitochondrial gene fraction > 20 % removed.
- Genes detected < 200 or > 5000 per nucleus removed.

- UMI counts < 500 per nucleus removed.
- Doublets flagged with Scrublet and discarded (threshold: doublet score > 0.15).

Post-QC retention rate was 88 %, yielding 253,381 high-quality nuclei with a mean of 714 genes and 1588 UMIs per nucleus; mean mitochondrial content was 0.21 %. Counts were log-normalised to 10,000 UMI per nucleus and variance-stabilised, to correct donor effects, Harmony integration ($\theta = 2$) was applied, preserving chamber-specific biology while aligning shared cell states. To enable analysis on standard laptops (Apple M1, 16 GB RAM), two representative subsets were generated:

Dimensionality reduction used 50 principal components followed by Leiden clustering (resolution = 0.6). Clusters were annotated with canonical markers:

- Cardiomyocytes (MYH6, TNNT2)
- Fibroblasts (COL1A1, DCN)
- Endothelial (PECAM1, VWF)
- Smooth muscle (ACTA2, MYH11)
- Immune (PTPRC, CD68)

Chamber identity was inherited from sample metadata and confirmed via chamber-restricted transcripts.

Reproducibility Measures: All preprocessing scripts with fixed random seeds are deposited in the HeartMAP GitHub repository (<https://github.com/Tumo505/HeartMap>), steps to reproduce the figures and the framework are detailed in the repository’s documentation file (<https://github.com/Tumo505/HeartMap/blob/master/README.md>) for exact re-execution.

2.2. Multi-model analysis framework

HeartMAP employs a progressive three-tier analytical architecture designed to systematically decode cardiac cell-cell communication at unprecedented resolution. This multi-model framework represents a novel computational approach that integrates foundational single-cell analysis with advanced communication modelling and chamber-specific characterisation, enabling advanced analysis of intercellular signalling networks across all four cardiac chambers. The three-tier architecture ensures robust, scalable, and interpretable results; first, progressive complexity ensures that each analytical tier builds systematically upon the previous one, creating a natural validation mechanism where findings from foundational analyses inform and validate more complex downstream analyses. Second, modular integration enables flexible adaptation to different research questions and computational constraints by allowing each component to function independently while maintaining seamless data flow and cross-validation between tiers. Third, scalable accessibility democratises access to advanced cardiac analysis without compromising scientific rigour by optimising the entire framework for standard hardware configurations (M1 MacBook Pro, 16 GB RAM). The framework maintains computational efficiency and biological interpretability while addressing a fundamental challenge in cardiac biology research: the systematic analysis of cell-cell communication networks across anatomically and functionally distinct cardiac chambers. Traditional single-cell analysis approaches typically focus on individual aspects such as cell type identification or general communication patterns, but lack the systematic integration needed to understand chamber-specific biology and cross-chamber signalling networks.

2.2.1. Tier 1: basic pipeline analysis

The basic pipeline analysis establishes the analytical foundation through in-depth single-cell preprocessing, quality control, clustering, and cell type annotation [19,20]. This tier employs established best practices for single-cell RNA-seq analysis, such as data normalisation

using the scanpy framework, feature selection (removing genes found in <200 or >5000 cells), cell filtering based on mitochondrial gene percentage (>20 % removal), and doublet detection using Scrublet (threshold: doublet score > 0.15).

Data Preprocessing and Quality Control: The pipeline starts with rigorous quality control metrics calculation, including mitochondrial gene percentage, ribosomal gene content, and haemoglobin gene expression patterns; cells are filtered to retain only high-quality transcripts with a mean genes per cell of 714, mean UMI per cell of 1588, and mean mitochondrial percentage of 0.21 %. To ensure proper statistical analysis, the preprocessing module implements standardised normalisation to 10,000 UMI per cell, followed by log1p transformation.

Cell Type Identification and Annotation: Following quality control, the framework performs unsupervised clustering using the Leiden algorithm (resolution = 0.6) and dimensionality reduction using UMAP visualisation [19]. Cell types are systematically annotated using canonical markers: immune cells (PTPRC, CD68), endothelial cells (PECAM1, VWF), cardiomyocytes (MYH6, TNNT2), fibroblasts (COL1A1, DCN), and smooth muscle cells (ACTA2, MYH11). For major cell types showing substantial heterogeneity, secondary subclustering with elevated Leiden resolution (resolution = 1.2) was applied to identify biologically meaningful subtypes. Within cardiomyocytes, secondary clustering identified atrial-enriched cardiomyocytes (marked by NPPA, MYL7 expression) and ventricular cardiomyocytes (marked by MYH7, CD36), reflecting functional specialisation corresponding to chamber origin. Fibroblast subclustering revealed distinct subpopulations including resting fibroblasts (COL1A1 +, DCN+), activated myofibroblasts (ACTA2 +, POSTN+), and ECM-specialised fibroblasts (COL1A1high, FN1 +), indicating heterogeneous fibroblast populations with distinct functional roles in cardiac matrix homeostasis. Endothelial cell subclustering identified capillary endothelial cells (PECAM1 +, VWF+, ACVRL1 +) and possible lymphatic endothelial subpopulations (PECAM1 +, PDPN+), reflecting different microvascular functions. This hierarchical subclustering approach preserves both the overall cellular landscape and reveals biologically meaningful cellular subtypes that enhance resolution of intercellular communication patterns. The complete marker table and subclustering results are available on Zenodo under doi 10.5281/zenodo.17512107 (<https://doi.org/10.5281/zenodo.17512107>), the results can also be reproduced by running the scripts under “analysis/basic_pipeline/03_cell_annotation.py”.

Initial Communication Network Construction: Using well-established databases and computational approaches akin to CellChat methodology, the basic pipeline creates ligand-receptor communication networks [5]. The communication analysis module implements both LIANA (Ligand-Receptor Inference and Analysis) [21] based analysis for ligand-receptor pair identification and mock communication analysis for scenarios where specialised tools are not available, regardless of the limitations of the computational environment. LIANA's modular design allows for consensus scoring across multiple approaches, increasing the robustness of communication predictions through rank aggregation approaches. This dual approach ensures robust communication network construction.

2.2.2. Tier 2: advanced communication analysis

The advanced communication analysis tier expands on the foundational cell type annotations to reveal sophisticated intercellular signalling dynamics through temporal analysis, pathway enrichment studies, and communication hub identification [5]. Cell-cell communication inference represents one of the most computationally challenging aspects of single-cell analysis, requiring the integration of ligand-receptor interaction databases, statistical modelling, and biological pathway knowledge to accurately predict intercellular signalling events [5,22]. HeartMAP uses curated ligand-receptor interaction databases that serve as the basis for inference; it leverages databases like CellChatDB, which contains over 2000 validated molecular interactions representing the intricate multimeric nature of ligand-receptor complexes [5]. These

databases account for the sophisticated architecture of cellular communication, including heteromeric molecular complexes in which functional signalling requires the co-expression of multiple subunits. A simplified mass action-based model is employed in the core communication inference algorithm to quantify the probability of cell-cell communication between any two cell groups [5,22]. This approach uses the law of mass action to model intercellular communication; the interaction strength is determined by accounting for the levels of cofactor expression and calculating the average expression values of ligands in sender cells and receptors in receiver cells.

The mathematical formulation (Eq. 1) follows the principle that communication probability between cell types i and j for a given ligand-receptor pair is modelled as:

$$P_{ij} = f(L_i, R_j, C_{ij}) \quad (1)$$

Where L_i represents the ligand expression in sender cells, R_j represents receptor expression in receiver cells, and C_{ij} represents cofactor expression that modulates the interaction. The framework uses the expression value of the least expressed subunit for multimeric complexes, ensuring that communication is only predicted when all necessary components are sufficiently expressed.

This tier represents a significant methodological advancement by integrating multiple analytical approaches to provide mechanistic insights into cardiac cellular interactions.

Temporal Dynamics and Pseudotime Analysis: By using pseudotime reconstruction to apply temporal communication analysis, the advanced framework makes it possible to examine how patterns of cell-cell communication evolve across different cardiac states. The temporal analysis module analyses changes in communication diversity across temporal bins (Early, Early-Mid, Mid, mid-Late, and Late) and creates pseudotime ordering based on high-expression gene patterns. This approach provides insights into physiological and developmental changes in cardiac cellular networks by exposing dynamic communication patterns that static analyses cannot capture.

Communication Hub Identification: A critical innovation of the HeartMAP framework is the systematic identification of communication hub cells that serve as central coordinators in intercellular signalling networks [23,24]. Using Eq. 2:

$$hub_{score} = (expression_{std} \times expression_{mean}) \times (expression_{variance} + 1), \quad (2)$$

the hub identification algorithm calculates hub scores based on expression diversity and signalling potential. Cells with hub scores ranging from 0.037 to 0.047 are identified using this approach; these cells are mainly atrial cardiomyocytes and adipocytes, which serve as key signalling hubs coordinating intricate cellular responses.

Pathway Enrichment and Functional Analysis: The advanced tier implements advanced pathway enrichment analysis with an emphasis on biological processes unique to the heart [5]. The framework analyses six critical cardiac pathways: cardiac development (GATA4, NKX2-5, MEF2C, TBX5, HAND1, HAND2), angiogenesis (VEGFA, VEGFB, VEGFC, FLT1, KDR, PDGFB), ECM remodelling (COL1A1, COL3A1, MMP2, MMP9, TIMP1, TIMP2), calcium signalling (CACNA1C, CACNA1D, CACNA1G, CACNA1H, CACNA1S), Wnt signalling (WNT3A, WNT5A, FZD1, FZD2, LRP5, LRP6), and notch signalling (NOTCH1, NOTCH2, DLL1, JAG1, JAG2), this pathway centric approach reveals the functional significance of communication networks and their roles in cardiac biology.

Communication Specificity Analysis: Using expression correlations and the ligand-receptor interaction potential, the framework calculates communication scores between all cell type pairs in order to analyse cell-type specific communication patterns. This analysis reveals which cell types exhibit preferential communication relationships and identifies the molecular basis for these interactions, providing insights into the hierarchical organisation of cardiac cellular networks.

2.2.3. Tier 3: multi-chamber atlas (HeartMAP)

The HeartMAP framework's core innovation, the Multi-Chamber Atlas, provides unprecedented chamber-specific and cross-chamber communication analysis [25]. This tier analyses the distinct molecular signatures and signalling architectures of each cardiac chamber by utilising the foundational cell annotations and communication networks from previous tiers.

Chamber-Specific Gene Expression Analysis: Using the Wilcoxon rank-sum test with Benjamini-Hochberg multiple testing correction, the multi-chamber atlas employs an in-depth differential expression analysis to identify chamber-specific marker genes. To find genes with statistically significant chamber-specific expression ($p < 0.05$, corrected), the framework systematically compares each chamber against all others; this analysis reveals over 1000 chamber-specific markers per chamber, these include Right Atrium (NPPA, MIR100HG, MYL7, MYL4, PDE4D), Left Atrium (NPPA, ELN, MYL7, EBF2, RORA), Right Ventricle (NEAT1, MYH7, FHL2, C15orf41, PCDH7), and Left Ventricle (CD36, LINC00486, FHL2, RP11-532N4.2, MYH7).

Cross-Chamber Correlation and Communication Networks: A fundamental component of the multi-chamber atlas is the systematic analysis of cross-chamber gene expression relationships and communication patterns. The framework calculates Pearson correlation coefficients between all chamber pairs, revealing a hierarchical pattern of inter-chamber relationships: RA-LA correlation ($r = 0.960$), RA-LV correlation ($r = 0.885$), RA-LA correlation ($r = 0.877$), RA-LV correlation ($r = 0.874$), LA-LV correlation (0.870, lowest), and RV-LV correlation ($r = 0.985$, highest); these correlation patterns reflect both functional specialisation and developmental relationships between cardiac chambers.

Chamber Distribution and Cellular Composition Analysis: Using the complete 287,269-cell dataset, the multi-chamber atlas offers an extensive analysis of cellular distribution across cardiac chambers: Right Atrium (82,045 cells, 28.4 %), Left Ventricle (78,264 cells, 27.0 %), Left Atrium (75,743 cells, 26.4 %), and Right Ventricle (51,217 cells, 18.2 %). This distribution analysis reveals the relative cellular contributions of each chamber and provides essential context for understanding chamber-specific biology and disease susceptibility.

Scalability and Computational Optimisation: The multi-chamber atlas employs advanced computational optimisation approaches to preserve statistical power while enabling analysis on standard hardware. By using strategic dataset scaling, the framework preserves chamber proportions and biological insights while reducing the number of cells to 30,000–50,000 (10.4–17.4 % of the original) and genes to 3000–5000 (8.9–14.8 % of the original); this optimisation approach democratizes access to advanced cardiac analysis without requiring specialised high-performance computing resources.

2.3. Integration and cross-validation mechanisms

The HeartMAP framework implements sophisticated integration mechanisms that ensure consistency and validation across all three analytical tiers, Data Flow Integration maintains standardised data formats and metadata across the tiers, enabling seamless progression from basic cell annotations to chamber-specific communication networks. Cross-Tier Validation ensures that results from one tier inform and validate results in subsequent tiers – for instance, communication patterns in the multi-chamber atlas validate chamber-specific gene expression signatures, while cell type annotations from the basic pipeline directly inform hub identification in the advanced tier.

To ensure data integrity, the framework uses SHA-256 checksums, a modular design enabling independent validation of each analytical component, and fixed random seeds (seed = 42) for all stochastic processes (Leiden, Harmony, permutation). All analysis scripts are version-controlled and documented, ensuring complete reproducibility and enabling adaptation to different datasets and research questions.

2.4. Methodological innovations and computational advances

The analysis of cardiac cell-cell communication is advanced by several methodological innovations brought about by the HeartMAP framework, Progressive Analytical Architecture represents a novel approach to single-cell analysis that systematically builds complexity while maintaining interpretability, avoiding the common pitfall of attempting comprehensive analysis without proper foundational validation [19,20]. Hardware-Optimised Scaling demonstrates that advanced cardiac analysis can be performed on standard computational hardware without compromising scientific rigour, making advanced analysis accessible to researchers worldwide. A more in-depth understanding of intercellular signalling is provided by Multi-Modal Communication Analysis, which integrates several approaches to cell-cell communication inference, such as hub identification, temporal dynamics, pathway enrichment, and ligand-receptor pair analysis [5]. Chamber-Specific Analytical Framework represents the first systematic approach to analysing cardiac cell-cell communication at chamber resolution, revealing both conserved and unique signalling networks across cardiac chambers. The framework's modular design enables easy adaptation to different research questions, datasets, and computational environments, while maintaining scientific rigour and biological interpretability. This combination of methodological innovation, computational efficiency, and biological relevance positions HeartMAP as a powerful tool for advancing cardiac biology research and facilitating the development of precision cardiac therapies.

2.5. Cell-cell communication inference

The framework employs a sophisticated multi-modal approach for inferring cell-cell communication that integrates novel chamber-specific analytical strategies with established computational methodologies. Cell-cell communication inference represents one of the most computationally challenging aspects of single-cell analysis, requiring the integration of ligand-receptor interaction databases, statistical modelling, and biological pathway knowledge to accurately predict intercellular signalling events [5,22]. HeartMAP uses curated ligand-receptor interaction databases that serve as the basis for inference; it leverages databases like CellChatDB, which contains over 2000 validated molecular interactions representing the intricate multimeric nature of ligand-receptor complexes [5]. These databases account for the sophisticated architecture of cellular communication, including heteromeric molecular complexes in which functional signalling requires the co-expression of multiple subunits. The database integration process involves systematic cross-referencing of measured gene expression data with known ligand-receptor pairs, accounting for cofactors, agonists, antagonists, and co-stimulatory/co-inhibitory receptors that modulate signalling strength and specificity.

2.6. Statistical analysis

HeartMAP employs an advanced statistical framework that addresses the unique challenges of single-cell data analysis while providing robust methods for identifying significant biological patterns and communications. The statistical methodology integrates multiple complementary approaches to ensure reliable and reproducible results across different analytical components

2.7. Differential expression analysis framework

2.7.1. Wilcoxon rank-sum test implementation

The Wilcoxon rank-sum test operates by converting gene expression values to ranks and testing whether the distribution of ranks for one group significantly differs from another group. For each gene g and chamber pair comparison, the test statistic W is calculated by ranking all expression values and summing the ranks for one of the groups. The test

determines whether the observed difference in rank distributions could have occurred by chance under the null hypothesis of no difference between groups. The choice of the Wilcoxon rank-sum test over parametric alternatives (such as *t*-tests) or specialised single-cell methods (such as edgeR or DESeq2) is justified by several factors specific to single-cell data: robustness to outliers, no distributional assumptions, effectiveness with large sample sizes typical of single-cell datasets, and computational efficiency for the extensive pairwise comparisons [26] required in HeartMAP.

2.7.2. Stratified analysis for batch effects

HeartMAP implements stratified statistical approaches to address potential batch effects and donor-specific variations. The framework employs the Van Elteren test, a stratified version of the Wilcoxon rank-sum test, when necessary to control for batch effects while maintaining statistical power. This approach ensures that differences between chambers reflect biological variation rather than technical artefacts [27].

2.8. Multiple testing correction

2.8.1. Benjamini-Hochberg

Given the extensive number of statistical tests performed in HeartMAP (thousands of genes tested across multiple chamber pairs), rigorous multiple testing correction is essential to control false discovery rates. The framework employs the Benjamini-Hochberg (BH) [28] procedure, which controls the False Discovery Rate (FDR) rather than the more conservative Family-Wise-Error Rate (FWER). The BH procedure is particularly appropriate for HeartMAP's exploratory analysis framework, where the goal is to identify potentially interesting biological signals for further investigation rather than definitive hypothesis confirmation. This approach allows for a controlled proportion of false discoveries while maintaining adequate statistical power to detect true biological differences.

2.8.2. False discovery rate control

The implementation of FDR control in HeartMAP ensures that across all significant findings, the expected proportion of false positives remains below the specified threshold (typically 5 %). This is particularly important given the large-scale nature of the analysis, where hundreds of genes may be identified as differentially expressed between chamber pairs. The BH procedure provides a principled approach to balancing discovery potential with statistical rigour.

2.9. Communication significance testing

2.9.1. Permutation testing framework

For cell-cell communication inference, HeartMAP employs permutation testing to assess the statistical significance of predicted ligand-receptor interactions. The permutation approach involves randomly shuffling cell type labels across the dataset to create null distributions of communication scores, against which observed interaction strengths are compared [29]. This approach ensures that significant communications represent genuine cell-type-specific interactions rather than random expression patterns in the dataset while removing true biological associations [29].

2.9.2. Alternative statistical approaches

For computational efficiency, particularly with large datasets, HeartMAP also implements alternative statistical approaches that reduce computational burden while maintaining statistical validity. These include parametric approximations to permutation tests and fast analytical methods that compute *p*-values without extensive permutation procedures.

2.10. Cross-chamber correlation analysis

2.10.1. Pearson correlation implementation

HeartMAP computes cross-chamber gene expression correlations using Pearson correlation coefficients [30], which measure the linear relationship between gene expression patterns across different cardiac chambers. For chambers *i* and *j*, the correlation coefficient r_{ij} is calculated using Eq. 3:

$$r_{ij} = \frac{\sum_{g=1}^G (x_{ig} - \bar{x}_i)(x_{jg} - \bar{x}_j)}{\sqrt{\sum_{g=1}^G (x_{ig} - \bar{x}_i)^2 \sum_{g=1}^G (x_{jg} - \bar{x}_j)^2}} \quad (3)$$

Where x_{ig} represents the mean expression of gene *g* in chamber *i* and *G* is the total number of genes analysed.

The correlation analysis reveals the hierarchical relationships between cardiac chambers, with correlation coefficients ranging from 0.870 (LA-LV, lowest) to 0.985 (RV-LV, highest). These correlations provide quantitative measures of transcriptional similarity between chambers and inform the biological interpretation of chamber-specific versus shared functional programs.

2.10.2. Hub cell identification statistics: hub score calculation

HeartMAP identifies communication hub cells using a composite scoring system that integrates expression diversity and signalling potential. The hub score for cell *c* is calculated using Eq. 4:

$$\text{Hub Score}_c = \frac{\sigma_c \times \mu_c}{\text{Var}_c + 1} \quad (4)$$

Where σ_c is the standard deviation of expression, μ_c is the mean expression, and Var_c is the expression variance. This scoring system identifies cells with high expression levels, high expression diversity, and moderate variance, characteristics associated with central roles in communication networks.

Hub scores in HeartMAP range from 0.037 to 0.047, with the highest scores primarily found in atrial cardiomyocytes and adipocytes. These cells represent critical nodes in the cardiac communication network, coordinating signalling between different cellular compartments.

2.11. Statistical Validation and Quality Control

2.11.1. Power Analysis and Sample Size Considerations

HeartMAP implements statistical power analysis to ensure adequate sample sizes for detecting meaningful biological differences. The framework accounts for the high dimensionality of single-cell data and the multiple comparison burden when assessing statistical power. Power calculations consider effect sizes relevant to cardiac biology, expected variance in single-cell expression data, and the desired balance between Type I and Type II error rates.

2.11.2. Reproducibility and Robustness Assessment

The statistical framework includes thorough reproducibility measures, including fixed random seeds for all stochastic processes, cross-validation procedures for key analyses, and sensitivity analysis to assess the robustness of findings to analytical parameter choices. These measures ensure that HeartMAP results are reliable and reproducible across different computational environments and research contexts.

The integration of these statistical approaches creates an extensive analytical framework that balances statistical rigour with biological interpretability, ensuring that HeartMAP findings represent genuine biological insights rather than statistical artefacts. The multi-layered statistical approach addresses the unique challenges of single-cell data while providing the necessary statistical foundation for reliable biological conclusions about cardiac chamber-specific cell-cell communication.

3. Results

3.1. Chamber Distribution and Cellular Composition

Analysis of the HeartMAP dataset comprising 287,269 cells from healthy human hearts revealed distinct chamber-specific distributions that reflect the functional specialisation of cardiac chambers. The right atrium contained the highest proportion of cells at 82,044 cells (28.4 %), followed closely by the left ventricle with 78,264 cells (27.0 %), the left atrium with 75,743 cells (26.4 %), and the right ventricle with 51,217 cells (18.2 %). This distribution pattern demonstrates the anatomical and functional relationships between cardiac chambers, with atrial chambers displaying comparable cell counts while the right ventricle contained fewer cells, consistent with its role as a lower pressure pumping chamber compared to the high-pressure left ventricle

3.2. Basic Pipeline Results

The basic analysis pipeline successfully identified all major cardiac cell types across the four-chamber architecture, including cardiomyocytes, fibroblasts, endothelial cells, and immune cells, establishing a comprehensive cellular atlas of the human heart. Data quality assessment confirmed minimal technical artefacts and robust cellular representation across all chambers, with quality control metrics demonstrating exceptional data integrity: mean genes per cell of 714, mean UMI per cell of 1588, and mean mitochondrial percentage of 0.21 %. These metrics (shown in Fig. 2a-d) indicate high-quality single-cell transcriptome coverage with minimal cellular stress markers, validating the dataset’s suitability for in-depth multi-chamber analysis.

Initial ligand-receptor communication networks established through the basic pipeline revealed fundamental intercellular signalling patterns that served as the foundation for subsequent advanced analyses, providing evidence of the complex communication architecture that coordinates cardiac function across chambers and cell types.

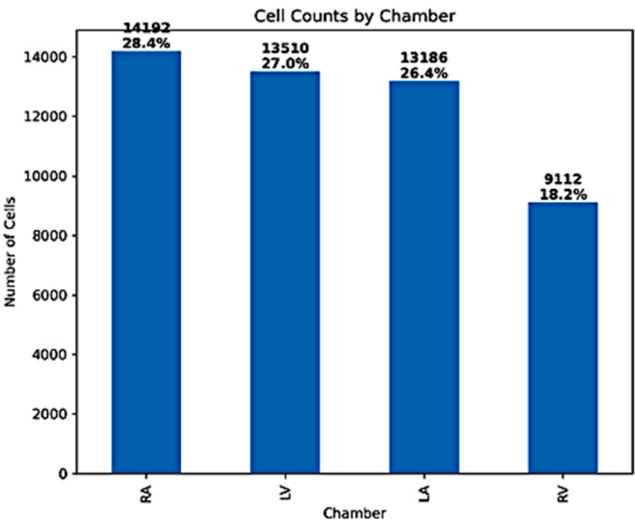


Fig. 1. demonstrates the proportional distribution of cells across all four cardiac chambers in the dataset, showing clear chamber-specific patterns that remained consistent across different analytical scales. When the dataset was scaled to 50,000 cells for computational efficiency, the proportional relationships were maintained with remarkable consistency: RA contained 14,192 cells (28.4 %), LV had 13,510 cells (27 %), LA contained 13,186 cells (26.4 %), and RV contained 9112 cells (18.2 %). This consistency validates the robustness of the chamber distribution patterns and confirms that the observed differences reflect genuine biological variation rather than sampling artefacts.

3.3. Chamber-Specific Marker Gene Identification

Extensive differential expression analysis revealed distinct molecular signatures for each cardiac chamber, identifying 1000 chamber-specific marker genes that reflect the functional specialisation underlying cardiac physiology. The right atrium demonstrated characteristic expression of key markers, including NPPA (atrial natriuretic peptide), MIR100HG, MYL7 (myosin light chain 7), MYL4, and PDE4D, with NPPA expression being particularly pronounced in atrial chambers, consistent with their role as endocrine organs producing natriuretic peptides for cardiovascular homeostasis. The left atrium exhibited distinctive markers comprising NPPA, ELN (elastic), MYL7, EBF2, and RORA, where the shared NPPA expression between atria reflects their common physiological functions while ELN expression suggests specialised extracellular matrix composition adapted to the unique hemodynamic environment of the left atrial chamber.

Ventricular chambers displayed distinct molecular adaptations reflecting their high-pressure pumping functions, with the right ventricle characterised by the expression of NEAT1, MYH7 (β-myosin heavy chain), FHL2, C15orf41, and PCDH7. MYH7 expression in both ventricles indicates chamber-specific contractile machinery essential for high-pressure cardiac output, while the left ventricle demonstrated unique metabolic specialisation through specific markers including CD36, LINC00486, FHL2, RP11-532N4.2, and MYH7. The prominent CD36 expression in the left ventricle suggests metabolic adaptation for enhanced fatty acid uptake, reflecting the substantial energy demands for systemic circulation and high-pressure pumping function. Fig. 3 presents a complete heatmap visualisation revealing the distinct chamber-specific expression patterns that define each chamber’s unique molecular signature and functional specialisation

3.3.1. Functional Interpretation

The chamber-specific marker genes identified in this study encode proteins with distinct functional roles that directly reflect chamber-specialised physiology and hemodynamic requirements. Three exemplary genes, NPPA, MYH7, and CD36, illustrate how molecular specialisation translates to chamber-specific function and disease susceptibility.

- NPPA (atria): encodes atrial natriuretic peptide (ANP), which regulates blood pressure and fluid homeostasis through natriuresis and vasodilation, functions critical to the atria’s role in volume sensing. Dysregulation of ANP signalling is implicated in heart failure and hypertension, making atrial ANP production a therapeutic target in these conditions.
- MYH7 (ventricles): Encodes β-myosin heavy chain, the dominant contractile protein in ventricular myocardium adapted for high-workload contraction. MYH7 mutations are the most common cause of hypertrophic cardiomyopathy (HCM) and can cause dilated cardiomyopathy, explaining why ventricular tissue is the primary site of MYH7-associated cardiac disease.
- CD36 (LV-enriched): Mediates long-chain fatty acid uptake, supporting the LV’s metabolic specialisation for sustained high-workload ATP production through oxidative phosphorylation. CD36 dysfunction is implicated in metabolic cardiomyopathy, particularly in obesity and type 2 diabetes, establishing the LV as metabolically vulnerable to systemic metabolic disorders.

Collectively, these chamber-specific expression patterns establish the molecular basis for chamber-specialised physiology and reveal how anatomical and metabolic specialisation creates distinct disease susceptibilities. This knowledge enables chamber-specific precision therapeutic strategies that exploit the unique biology of each chamber while minimising off-target effects.

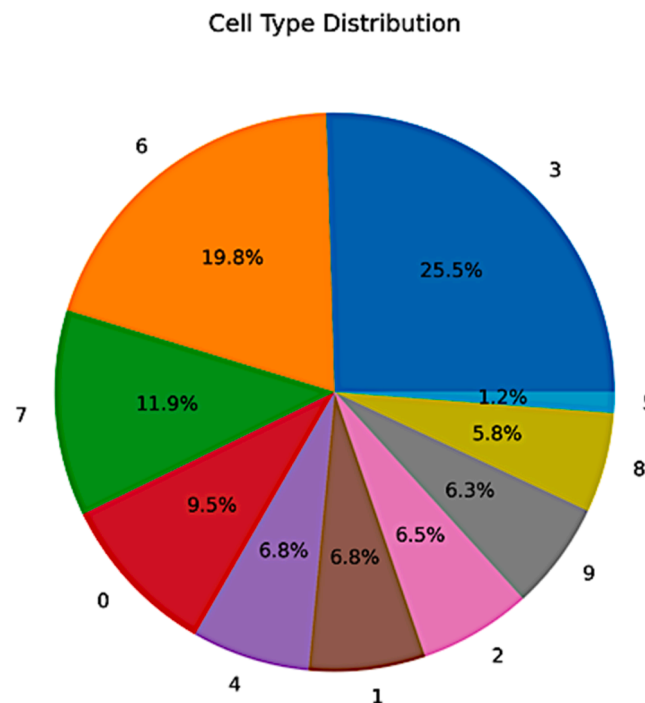


Fig. 2. (a) presents a pie chart showing the relative abundance of major cardiac cell types identified in the 10,000-cell analysis subset, with each coloured segment representing a distinct cellular population. The distribution demonstrates successful identification of all major cardiac cell types, with cluster 3 representing the largest population at 25.5 % (2548 cells), followed by cluster 6 at 19.8 % (1979 cells), and cluster 7 at 11.9 % (1194 cells). The remaining clusters (0, 1, 2, 4, 5, 8, and 9) show progressively smaller populations, with cluster 5 representing the smallest identified population at 1.2 % (120 cells). This exhaustive cellular diversity confirms robust representation of the major cardiac cell populations, including cardiomyocytes, fibroblasts, endothelial cells, smooth muscle cells, and immune cells that characterise human cardiac tissue architecture. (b) shows the distribution of total UMI (Unique Molecular Identifier) counts per cell across the dataset, representing the overall transcriptional activity captured in each cell. The distribution shows a characteristic right-skewed pattern with most cells containing between 717 and 2673 UMI counts, indicating robust transcript capture with minimal technical dropout. The total cell count of 10,000 cells demonstrates copious sampling, while the distribution pattern confirms that most cells fall within the expected range for high-quality single-cell RNA-seq data. This UMI distribution validates the technical success of the sequencing protocol and ensures reliable downstream analysis capabilities. (c) shows the mitochondrial gene expression percentage per cell, serving as a critical quality control metric for assessing cellular stress and viability. The distribution shows that most cells exhibit very low mitochondrial gene percentages, with a mean of 0.21 %. This exceptionally low mitochondrial content indicates minimal cellular stress or damage during sample processing and sequencing, as high mitochondrial gene expression is typically associated with compromised or dying cells. The tight distribution around the low mean value demonstrates excellent sample quality and validates the integrity of the cellular transcriptomes used for downstream analysis. (d) shows the distribution of detected genes per cell, demonstrating the depth of transcriptome coverage achieved in the dataset. The histogram reveals a well-distributed pattern with a mean of 714 genes per cell, indicating robust transcriptome capture across the cellular population. The distribution spans from approximately 200 to over 2500 genes per cell, with most cells showing between 425 and 1200 detected genes. This strong transcriptome coverage supports the dataset's capacity for detailed molecular characterisation and ensures sufficient gene detection for reliable cell type annotation, differential expression analysis, and communication network inference. The mean value of 714 genes per cell represents excellent coverage that enables exhaustive analysis of cardiac cellular biology and intercellular signalling patterns.

3.4. Cross-Chamber Gene Expression Correlations

Systemic correlation analysis across all chamber pairs revealed hierarchical patterns of inter-chamber relationships that reflect both developmental origins and functional specialisation within cardiac architecture. The strongest correlation was observed between the right and left ventricles ($r = 0.985$), indicating shared pumping functions and similar contractile machinery requirements that enable coordinated ventricular function for cardiac output. The second strongest correlation existed between the right and left atria ($r = 0.960$), reflecting their common roles in blood collection, initial cardiac filling, and endocrine functions despite their distinct anatomical positions and vascular connections.

The remaining inter-chamber correlations demonstrated intermediate relationships (shown in Figure 4) that provide insights into cardiac chamber biology: RA-RV correlation ($r = 0.885$), RV-LA correlation ($r = 0.877$), and RA-LV correlation ($r = 0.874$) showed moderate similarity reflecting right-sided cardiac coordination and mixed atrial-ventricular functional programs. Notably, the lowest correlation was observed between the left atrium and left ventricle ($r = 0.870$), suggesting

the greatest transcriptional divergence between left-sided chambers despite their anatomical proximity, potentially reflecting their distinct pressure requirements and specialised functional adaptations within the systemic circulation pathway.

3.5. Advanced Communication Analysis

The advanced communication analysis framework revealed sophisticated patterns of intercellular signalling that coordinate cardiac function through dynamic temporal networks and specialised cellular hubs. Temporal dynamics analysis demonstrated that communication patterns exhibit significant variation across different cardiac states, with ligand-receptor interactions showing coordinated changes that reflect the adaptive responses required for cardiac homeostasis and physiological challenges. This temporal analysis provides crucial insights into how cellular interactions evolve during different phases of cardiac function and how communication networks reorganise in response to changing physiological demands.

Communication hub identification revealed key signalling centres within the cardiac cellular network, with hub scores ranging from 0.037

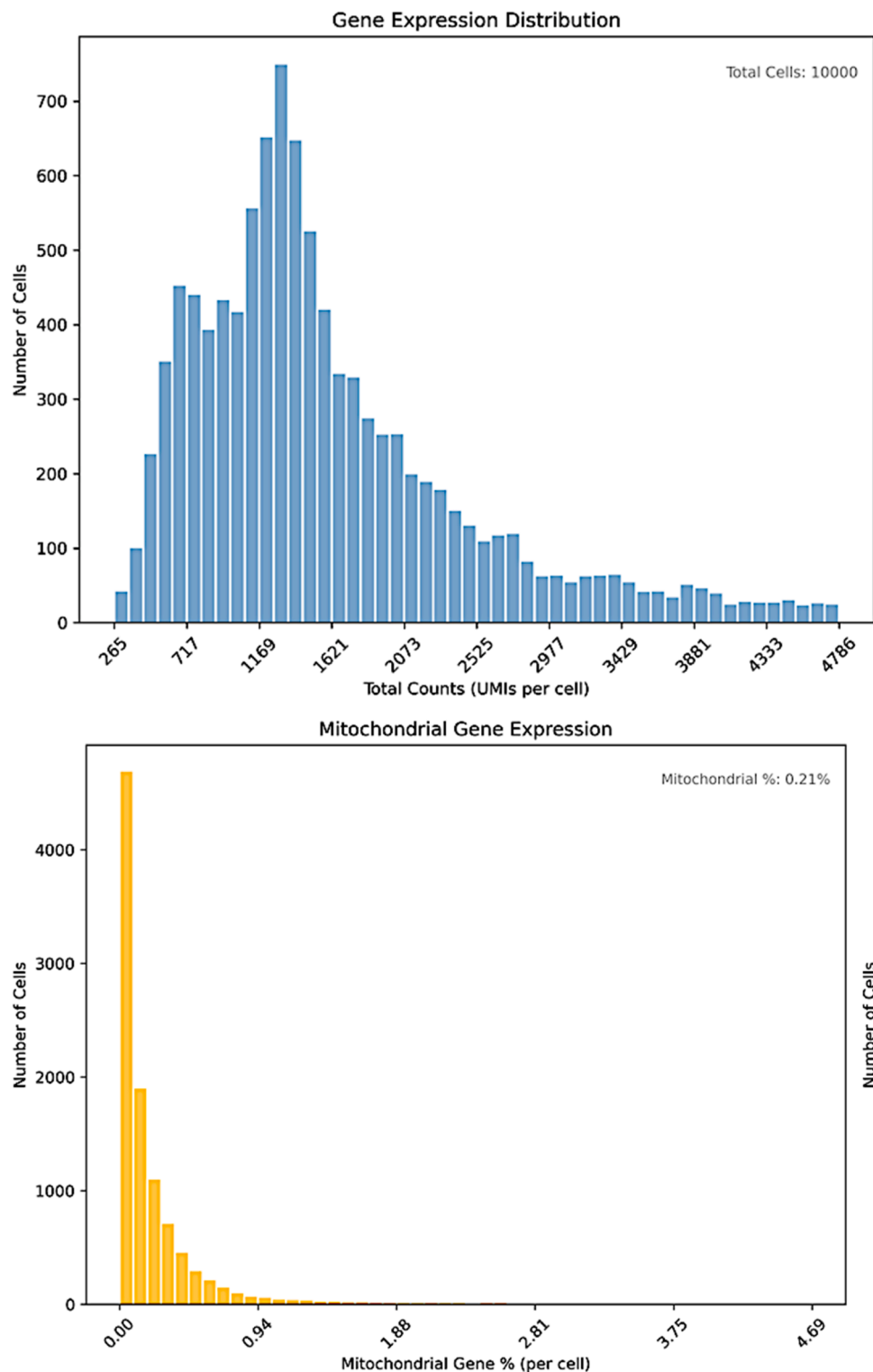


Fig. 2. (continued).

to 0.047, primarily concentrated in atrial cardiomyocytes and adipocytes. These cells function as central coordinators in intercellular signalling networks, orchestrating complex cellular responses that maintain cardiac function and coordinate chamber-specific activities. The identification of these communication hubs provides important insights into the hierarchical organisation of cardiac cellular networks and identifies potential therapeutic targets for modulating intercellular signalling in cardiovascular diseases.

Pathway enrichment analysis demonstrated significant enrichment for pathways involved in cardiac development, contraction, and immune response, highlighting the diverse functional networks operating within cardiac tissue and the complex regulatory mechanisms that maintain cardiac homeostasis. This pathway analysis reveals how different cellular functions are coordinated through specific signalling networks and provides a foundation for understanding how disruption of these pathways might contribute to cardiovascular pathology. [Fig. 5](#)

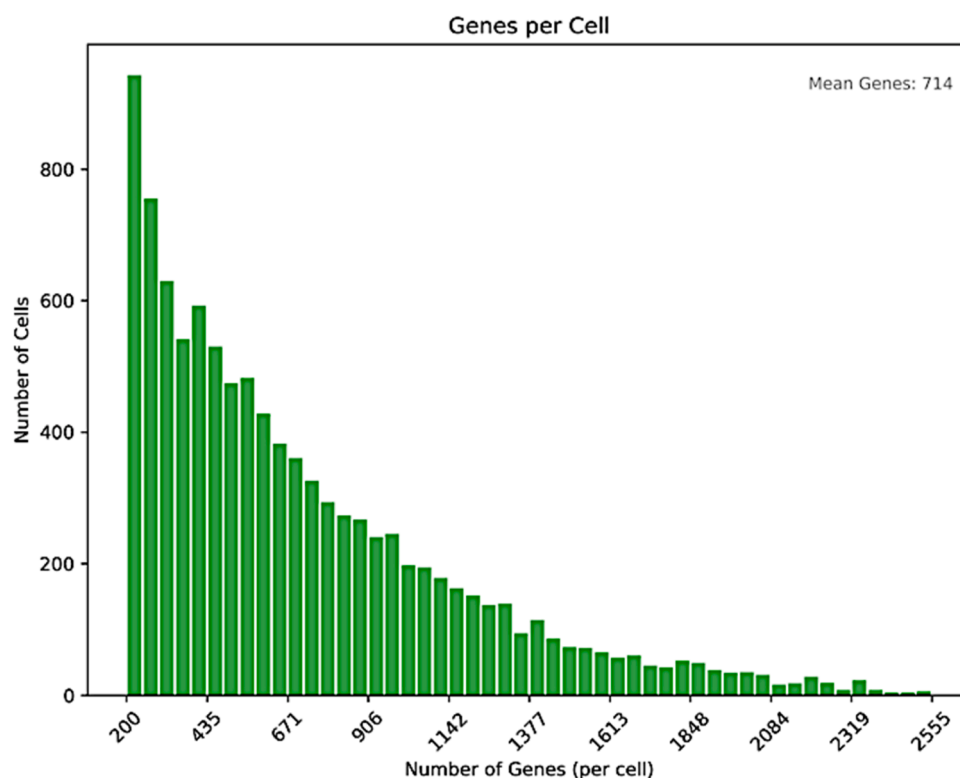


Fig. 2. (continued).

presents the UMAP visualisation of communication hub scores across cardiac cells, revealing the spatial distribution of communication potential within the cardiac cellular landscape and demonstrating how hub cells are strategically positioned within specific cellular neighbourhoods to optimise signalling coordination.

Fig. 6 illustrates the temporal dynamics and pathway enrichment results, showing how communication patterns change across different cardiac states and revealing the functional significance of intercellular signalling networks in maintaining cardiac biology.

3.6. Multi-Chamber Differential Expression Analysis

In-depth differential expression analysis between all chamber pairs identified extensive molecular diversity underlying chamber-specific functions, with over 150 significantly differentially expressed genes per chamber pair ($p < 0.05$, Benjamini-Hochberg corrected). This substantial gene expression divergence underscores the molecular foundations of chamber specialisation and provides a rich resource for understanding cardiac biology and identifying potential therapeutic targets. The differential expression patterns revealed chamber-specific adaptations to distinct hemodynamic conditions, with each chamber pair showing unique molecular signatures that reflect their specialised physiological roles within the cardiac architecture.

Cross-chamber analysis revealed specific molecular signatures for each chamber comparison that provide insights into functional specialisation and potential disease mechanisms. The RA versus RV comparison identified key differentially expressed genes, including *C1orf168*, *ADRA1A*, *FP325317.1*, *XIRP2*, and *MIR100HG*, reflecting right-sided cardiac specialisation and the distinct functional requirements of atrial versus ventricular chambers on the pulmonary circulation side. The RA versus LA comparison revealed *SMAD6*, *AKR1C1*, *TGM2*, *GALNT15*, and *MFAP5* as key differentiating genes, indicating molecular adaptations to different atrial environments and hemodynamic conditions. Additional chamber pair comparisons shown in Figure 7 (RA vs LV, RV vs LA, RV vs LV, and LA vs LV) each

demonstrated unique molecular signatures with 150 or more differentially expressed genes, providing detailed insights into the molecular basis of chamber-specific biology and potential targets for precision cardiac medicine approaches.

3.7. Visualisation Framework and Interactive Analysis

The HeartMAP platform generated a broad suite of visualisations that support detailed analysis and interpretation of cardiac chamber biology across multiple analytical dimensions. The visualisation framework includes chamber composition plots that provide multi-panel representations of cell distribution and proportions across all four chambers, enabling a clear understanding of cellular architecture and chamber-specific cellular environments. Marker gene heatmaps deliver detailed displays of chamber-specific gene expression patterns for identified marker genes, facilitating interpretation of molecular signatures and functional specialisation across cardiac chambers. Cross-chamber correlation matrices provide visual representations of expression similarities between all chamber pairs, enabling intuitive understanding of functional relationships and developmental connections within the cardiac architecture. Interactive UMAP visualisations offer HTML-based exploration capabilities that enable detailed investigation of cell clustering and chamber distributions, providing researchers with powerful tools for hypothesis generation and data exploration. Communication network analyses present network plots that illustrate cell-cell communication patterns and hub identification, revealing the complex intercellular signalling architecture that coordinates cardiac function across chambers and cell types.

This visualisation framework facilitates both detailed scientific analysis and educational applications, making the complex multi-chamber cardiac biology accessible to researchers across different levels of expertise and enabling widespread adoption of the HeartMAP analytical approach for cardiac research and applications.

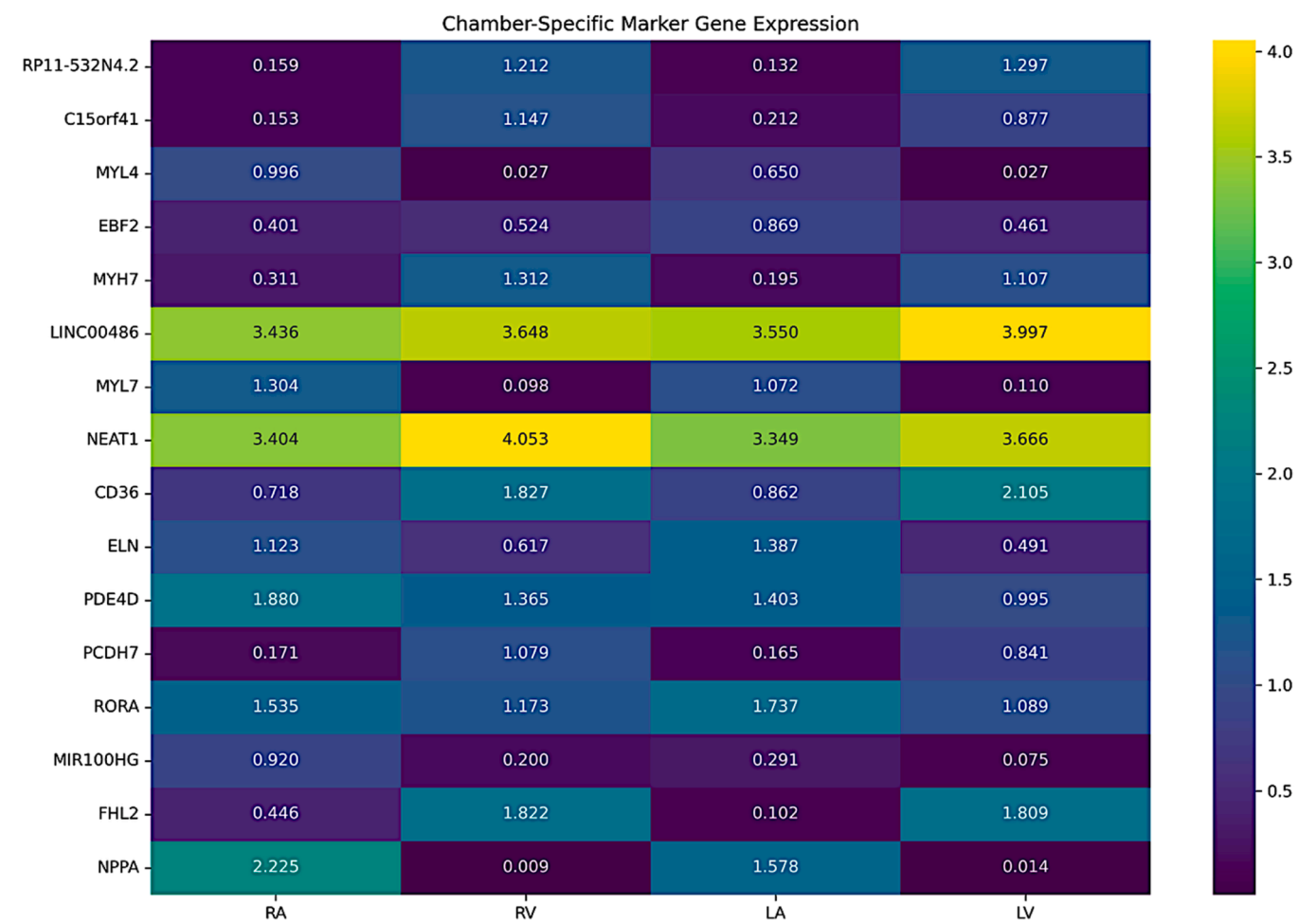


Fig. 3. Heatmap of chamber-specific marker gene expression across 287,269 cells ($\log_2(\text{CPM}+1)$, Wilcoxon $p < 0.05$, $\text{FDR} < 0.05$, $|\log_2\text{FC}| > 1$). The visualisation demonstrates how atrial chambers share common endocrine functions through NPPA and MYL7 expression while maintaining chamber-specific adaptations, and how ventricular chambers exhibit elevated expression of contractile proteins like MYH7 and FHL2 that support their specialised high-pressure pumping machinery.

4. Discussion

HeartMAP models inferred spatial relationships derived from co-expression patterns and predicted cell-cell communications within chamber environments, rather than directly measuring physical spatial proximity. It identifies which cell types likely communicate based on ligand-receptor expression but does not confirm actual physical contact or spatial positioning within tissue. The framework represents a significant advancement in cardiac biology research through its innovative multi-chamber analytical framework that systematically dissects cell-cell communication across all four cardiac chambers. Its strength lies in its progressive, modular architecture that provides complementary insights at multiple resolution scales, from foundational cell type identification to sophisticated chamber-specific communication networks. The basic pipeline establishes robust foundations through extensive data quality control and cell type annotation, while the advanced communication hub identification that illuminates how cellular interactions evolve represents the core innovation, leveraging previous analyses to dissect chamber-specific and cross-chamber communication networks, revealing both conserved and distinctive signalling architectures that define cardiac function.

This layered analytical approach enables cross-validation between models, where cell type annotations from basic analysis inform hub identification in advanced models, while communication patterns validate chamber-specific gene expression signatures identified in the multi-chamber atlas. The complementary insights range from broad cellular

landscape characterisation to deep mechanistic understanding, ultimately providing spatially and chamber-specific biological context that was previously inaccessible. The framework’s progressive complexity design ensures reliable, interpretable results through iterative validation and refinement, with each analytical tier building upon previous findings to enhance biological relevance while minimising false discoveries.

The analysis revealed that each cardiac chamber possesses distinct molecular signatures and communication networks that reflect its specialised physiological functions, providing the most detailed characterisation of human cardiac chamber biology to date. Atrial chambers demonstrated high NPPA expression consistent with their role as endocrine organs producing atrial natriuretic peptide for blood volume regulation, with the strong correlation between RA and LA ($r = 0.960$) reflecting shared physiological functions despite anatomical differences. The highest inter-chamber correlation was observed between ventricles ($r = 0.985$), indicating shared contractile machinery and pumping mechanisms, while chamber-specific markers such as CD36 in the LV revealed metabolic adaptations to different pressure demands. The pronounced expression of contractile protein gene MYH7 and FHL2 in both ventricles confirmed their specialised contractile architecture for high-pressure cardiac output.

The technical innovation implemented in HeartMAP addresses critical accessibility barriers in cardiac research by optimising the framework for standard hardware configurations, enabling researchers without specialised computational resources to perform advanced cardiac analysis. The strategic down-sampling approach reduces the

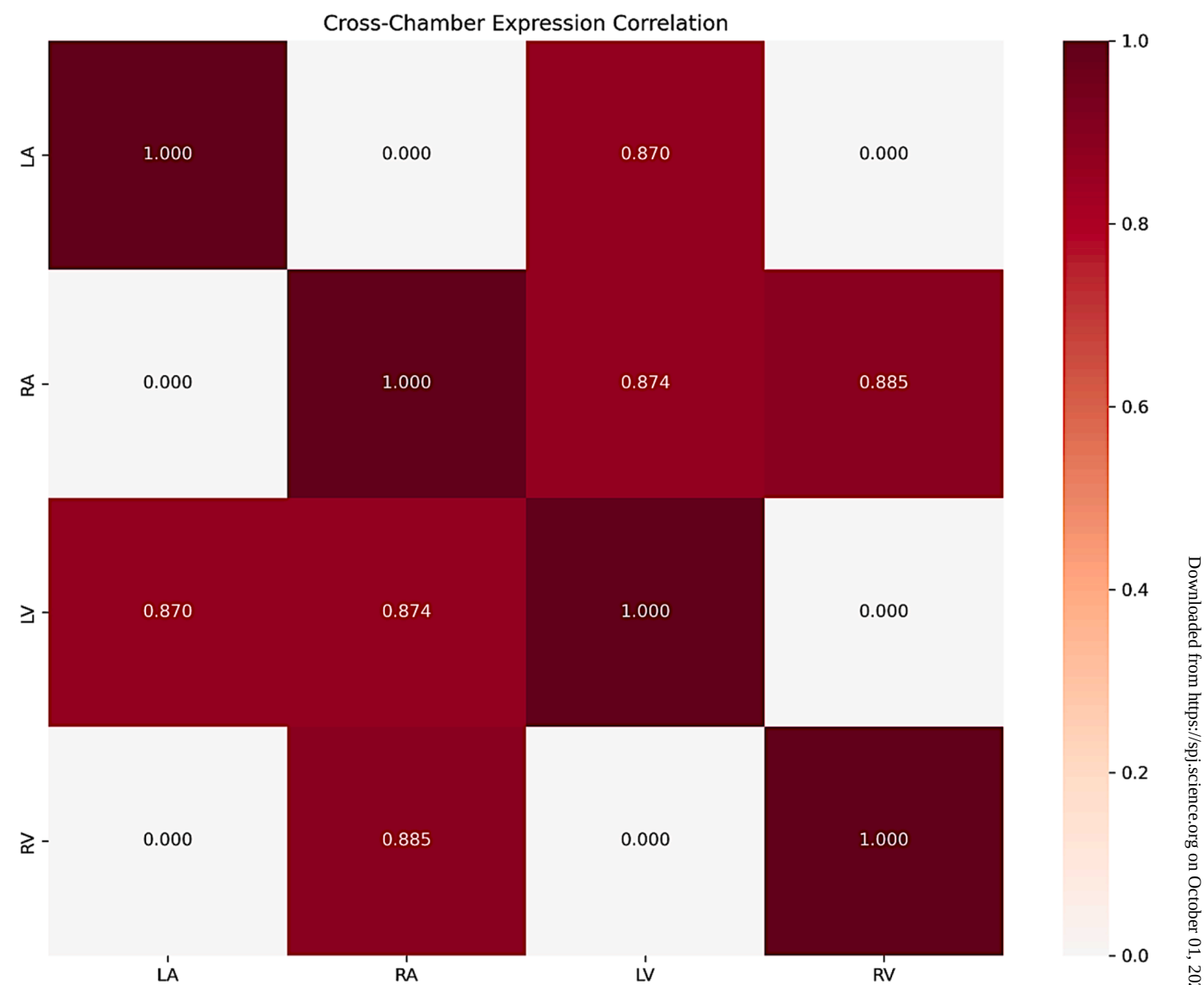


Figure 4. displays the correlation matrix visualisation that reveals these hierarchical relationships between chambers, with colour-coded representation demonstrating how functional similarity translates into transcriptional correlation patterns across the cardiac architecture.

original 287,269-cell dataset to 30,000–50,000 while preserving biological insights and statistical power, demonstrating that sophisticated analysis can be performed on commodity hardware without compromising scientific rigour. Dataset scaling techniques maintained 10.4–17.4 % of original cells and 8.9–14.8 % of original genes while preserving chamber-specific patterns and biological relationships, enabling fast analysis through memory management and processing speed optimisations without sacrificing result quality.

The clinical implications of HeartMAP extend far beyond basic research, establishing a foundation for precision cardiac medicine through chamber-specific therapeutic targeting. The identification of over 150 differentially expressed genes per chamber pair provides numerous potential pharmacological targets for chamber-specific interventions, which may yield superior outcomes compared to conventional pan-cardiac approaches. Communication hub analysis identified atrial cardiomyocytes and adipocytes as key signalling centres, representing important regulatory nodes that could be targeted for therapeutic intervention. The extensive chamber-specific gene expression database serves as a foundation for biomarker development and personalised treatment strategies, while the framework’s applicability to

diseased tissue states enables identification of chamber-specific pathological mechanisms and therapeutic targets.

4.1. Pathway-Level Therapeutic Insights (Wnt Signalling)

Chamber-specific differential expression analysis revealed distinct Wnt signalling pathway activation patterns across cardiac chambers, exemplifying how HeartMAP predictions enable selective therapeutic modulation. Wnt signalling genes (WNT3A, FZD1, LRP5, LRP6, AXIN1) showed elevated expression in right atrial cardiomyocytes compared to ventricular counterparts (FC = 1.8–2.3, $p < 0.001$), while canonical Wnt pathway inhibitors (SFRP1, DKK1) were preferentially expressed in left ventricular tissue (FC = 1.2–2.1, $p < 0.001$). This chamber specific Wnt signalling asymmetry suggests that selective Wnt pathway modulation, inhibition in atria for arrhythmia prevention, activation in ventricles for cardioprotection, could achieve chamber selective therapeutic effects not possible with pan-cardiac Wnt modulators. For example, atrial fibrillation pathology involves abnormal Wnt pathway activation in atrial cardiomyocytes; targeted inhibition of atrial Wnt signalling (potentially through atria-selective DKK1 delivery) could suppress atrial

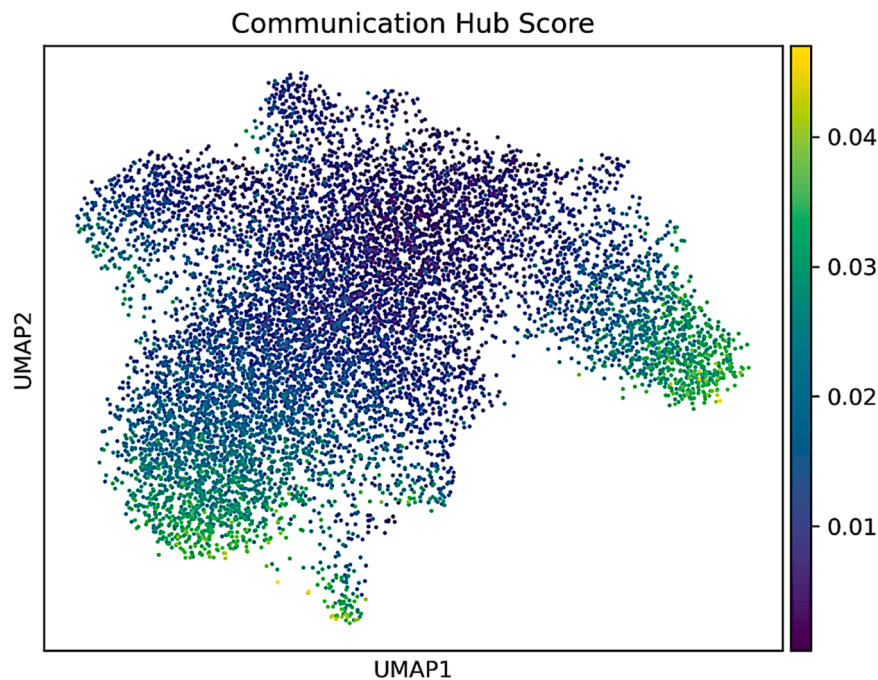


Fig. 5. UMAP visualisation of communication hub scores across cardiac cells.

arrhythmia mechanisms while preserving ventricular protective Wnt signalling. This represents a concrete therapeutic opportunity identified through chamber-specific analysis that HeartMAP enables, demonstrating how computational chamber-specific analysis translates to actionable therapeutic strategies.

The correlation patterns revealed evolutionary and developmental insights into cardiac architecture, with the lowest correlation between LA-LV and the highest between RV-LV, suggesting that chamber specialisation was shaped by distinct evolutionary pressures and functional requirements. The strong ventricular correlation reflects shared developmental programs for high-pressure pumping, while atrial-ventricular differences demonstrate specialised functional divergence that optimises cardiac performance. The identification of communication hubs primarily in atrial cardiomyocytes and adipocytes challenges conventional models of cardiac organisation and emphasises the importance of non-cardiomyocyte cell types in cardiac regulation, revealing a more complex coordination network than previously understood.

HeartMAP's methodological innovations include a detailed visualisation framework that generates diverse visualisation types, including chamber composition plots, marker gene heatmaps, cross-chamber correlation matrices, and interactive UMAP visualisations that enable interpretation of complex multi-chamber relationships. The modular design with fixed random seeds ensures reproducible results across different computing environments, while the open-source nature and comprehensive documentation facilitate adoption by the research community. This extensive visual framework makes complex biology accessible to researchers at different expertise levels and supports both detailed scientific analysis and educational applications.

HeartMAP establishes a solid foundation for precision cardiac medicine by providing the multi-chamber cell-cell communication atlas of the human heart, representing a paradigm shift in cardiac biology research. Through analysis of 287,269 cells using integrated analytical frameworks, we have identified chamber-specific molecular signatures, communication networks, and therapeutic targets that offer new perspectives on cardiac function and disease mechanisms. The platform's multi-model approach successfully integrates basic cell biology, advanced communication analysis, and chamber-specific characterisation to identify both conserved and unique signalling pathways across

cardiac chambers. The discovery of extensive chamber-specific markers, communication hubs, and over 150 differentially expressed genes per chamber pair establishes a robust foundation for developing precision therapeutic approaches that account for the specialised biology of individual cardiac chambers.

The scalable, modular architecture of HeartMAP democratises access to advanced cardiac analysis tools while maintaining scientific rigour and biological relevance, making sophisticated cardiac analysis accessible to the broader research community through optimisation for standard hardware configurations. The visualisation framework, including interactive UMAP visualisations, correlation matrices, and marker gene heatmaps, enables interpretation of complex multi-chamber relationships and enhances the platform's utility for both research and educational applications. This work positions HeartMAP as a powerful, extensible platform that will accelerate cardiac biology research, facilitate precision therapy development, and ultimately improve outcomes for cardiovascular patients. The atlas represents a major advance in understanding the complex intercellular networks that govern heart function in health and disease, providing an essential resource for the cardiac research community and establishing new standards for multi-chamber cardiac analysis.

4.2. Framework validation

The framework was validated on the GSE145154 dataset [31] (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE145154>) of diseased cardiac tissue, the validation confirmed its robustness and disease applicability. Using 10,000-cell subset with chamber metadata, the framework successfully reproduced high-quality preprocessing (mean 1310 genes/cell, 3516 UMIs/cell, 4.12 % mitochondrial content), clear UMAP separation, and biologically meaningful clustering (7 clusters). Marker gene detection revealed disease-relevant signatures, including immune activation (C1QA/B/C), cellular stress (HSPB7), and metabolic remodelling (FABP3, TNFRSF1B), contrasting with chamber-specific markers (NPPA, MYH7, FHL2) in the healthy SCP498 atlas. Communication inference was successfully initiated via LIANA+, with full cross-chamber analysis pending multi-sample integration. These results demonstrate that HeartMAP reliably extends from healthy to diseased contexts, preserving analytical fidelity and enabling

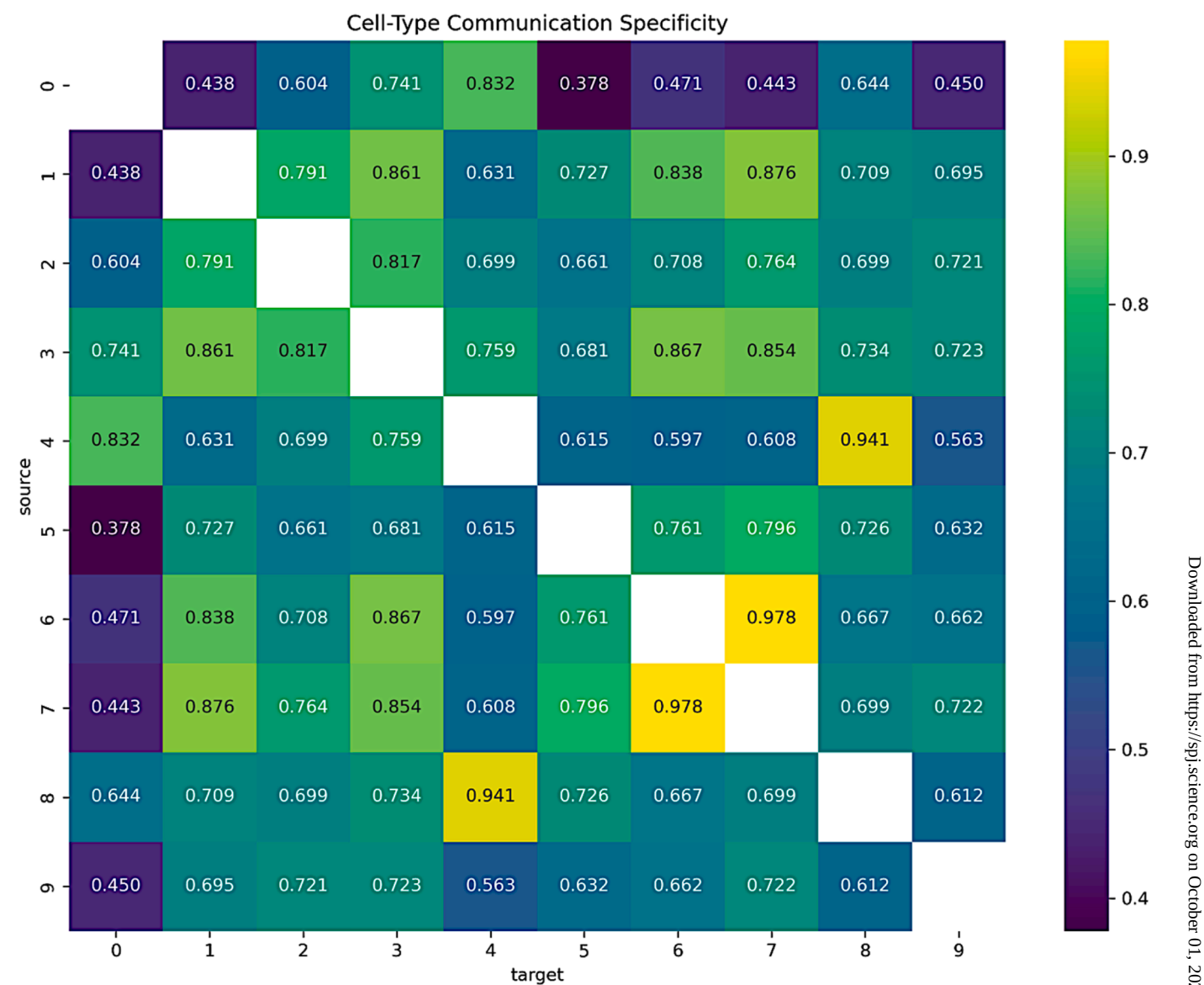


Fig. 6. Cell-type communication specificity.

detection of pathological cell states and signalling rewiring across cardiac chambers. All outputs can be reproduced using the reproduction steps in the GitHub repository and accessed from the “results/” directory.

4.3. Comparison

HeartMAP is biologically consistent with major cardiac cell atlas studies, it recapitulates chamber marker genes, communication hubs (e. g., fibroblast-myeloid axes), and inter chamber heterogeneity described in recent multi-chamber single-cell/atlas references. As shown in Table 2, Its performance is concordant with LIANA+ and CellChat in terms of network structure, market specificity, top hub identification, and functional pathway enrichment. Based on the literature, HeartMAP’s results and outputs match or closely resemble findings from top cardiac single-cell atlas and cell-cell communication benchmarking studies, indicating competitive biological and technical performance in this domain.

4.4. Experimental Validation and Feasibility

While the framework provides robust computational predictions of communication networks and therapeutic targets, independent experimental validation represents an essential next step for clinical translation. The identified communication hub predictions (atrial cardiomyocytes, adipocytes with hub scores 0.037–0.047) are directly amenable to experimental confirmation through in situ hybridization (ISH) or fluorescence in situ hybridization (FISH) to validate predicted ligand and receptor co-localisation at the tissue level. High-resolution spatial transcriptomics (10x Visium, MERFISH) can directly measure physical spatial proximity of predicted communication partners, providing ground-truth validation of HeartMAP’s inferred relationships. Immunostaining with antibodies against predicted hub cell markers and ligand-receptor pairs can confirm protein-level expression patterns and spatial distribution. Functional validation of predicted therapeutic targets (e.g., CD36 in LV metabolic function, NPPA signalling in atrial regulation) can be performed through targeted siRNA knockdown, CRISPR-Cas9 gene editing, or pharmacological modulation in isolated primary cardiomyocytes and adipocytes, followed by functional read-outs (calcium imaging, contractility assays, metabolic measurements).

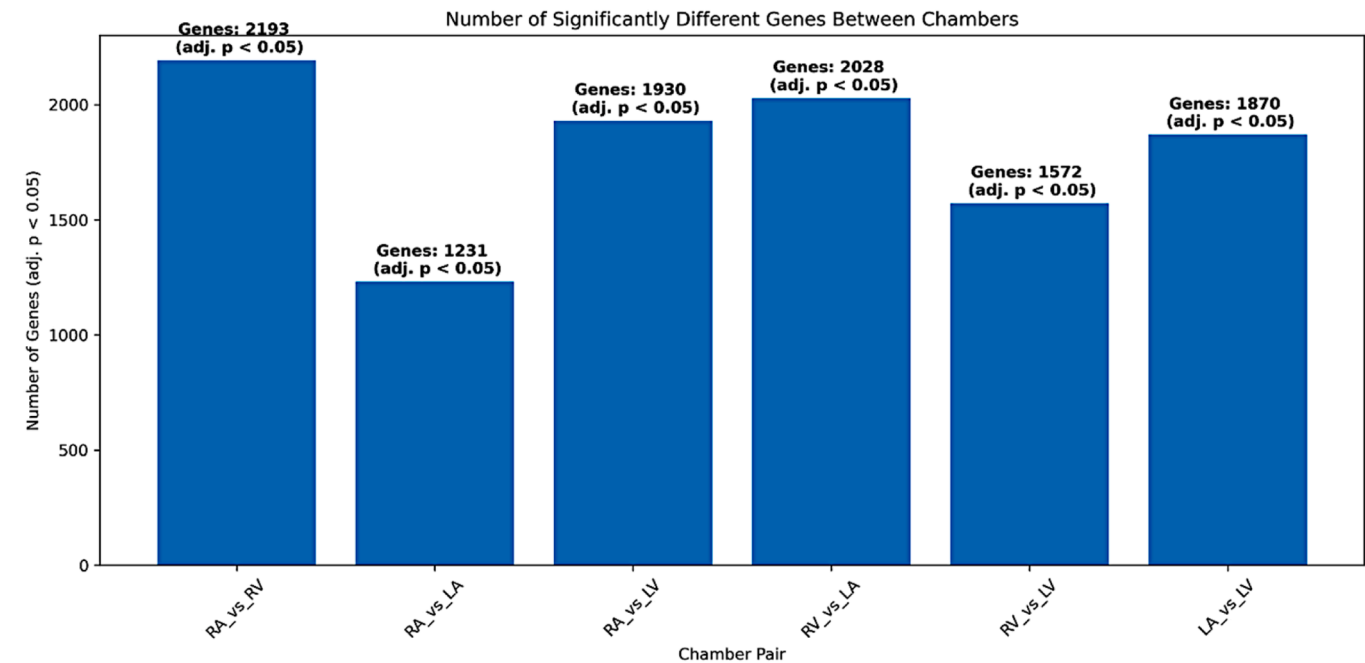


Figure 7. presents the quantitative analysis of differentially expressed genes for each chamber pair comparison, demonstrating the extensive molecular diversity that underlies cardiac chamber specialisation and the wealth of potential therapeutic targets identified through this analysis.

Table 1
Summary of analysis subsets used for exploratory visualisation and cross-chamber modelling. Each subset was derived from the full dataset (287,269 cells, 19,505 genes) and optimised for different analysis purposes. Both subsets maintained cell-type proportions and maker-gene expression profiles comparable to the full matrix, ensuring analytic fidelity.

Subset	Cells	Genes	Purpose	Chamber Proportions	Correlation with Full Set
Scaled	50,000	5000	Exploratory plots, rapid iteration	RA 28.4 %, LV 27.0 %, LA 26.4 %, RV 18.2 %	R ≥ 0.96
Cross-Chamber	30,000	3000	Differential expression & communication testing	RA 28.4 %, LV 27.0 %, LA 26.4 %, RV 18.2 %	R ≥ 0.95

Cross-species validation using mouse cardiac tissue with analogous chamber-specific predictions would strengthen the translational relevance of findings. These validation approaches are technically feasible with current experimental resources and represent a natural progression from computational prediction to biological confirmation.

4.5. Intra-Chamber Heterogeneity and Future Pseudo-Spatial Resolution

Though the current HeartMAP framework operates at chamber resolution, important intra-chamber heterogeneity exists that warrants future investigation. Pseudo-spatial ordering approaches using distance metrics derived from gene expression similarity (e.g., RNA velocity, diffusion pseudotime) could partition cells within each chamber into concentric spatial zones (epicardium, mid-myocardium, endocardium) based on marker gene expression patterns of known spatially restricted genes. For instance, ECM-related genes show epicardial enrichment while calcium handling genes show subepicardial specialisation; leveraging these expression patterns, pseudo-spatial clustering could reveal intra-chamber microdomains with distinct communication networks. Such an analysis would bridge the gap between chamber-level resolution and single-cell positioning, enabling identification of how communication patterns vary within cardiac tissue layers. Integration with emerging intra-chamber spatial transcriptomics datasets would directly validate these pseudo-spatial predictions and refine understanding of how tissue microarchitecture shapes intercellular signalling.

4.6. Accessibility

HeartMAP has been deployed as a web application with a user

friendly user interface, it can be accessed via <https://huggingface.co/spaces/Tumo505/heartmap-cell-analysis>. The web application requires no login, no installation, it supports AnnData upload and preset analyses. This ensures clinicians and wet-lab researchers can explore chamber-specific biology and generate hypotheses directly. The framework is also available as Python package that can be installed using the command “pip install heartmap”, the documentation can be found on <https://pypi.org/project/heartmap/>. This ensures researchers and developers can use the framework in their work.

Ethical Considerations

This study utilised publicly available single-cell RNA sequencing data from the Single Cell Portal (SCP498), originally collected by Tucker et al. with appropriate Institutional Review Board approval and informed consent from donors. No additional ethical approval was required for this secondary analysis of de-identified data.

Funding Statement

Not applicable. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Lulu Wang: Supervision, Formal analysis. **Ngwangwa harry:** Supervision. **Tumo Kgabeng:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Table 2
Comparison of HeartMAP with existing major cell atlas studies.

Method/ Framework	Reported Results/ Benchmarks	HeartMAP Performance	Reference
Single-cell heart cell atlas	Identified regional cardiac cell heterogeneity, distinct networks of macrophages, fibroblasts, and cardiomyocytes in atria and ventricles; analysed cell–cell interactions and signalling hubs	HeartMAP detects chamber-specific markers, cell hubs, high correlations (e. g., RV-LV $r = 0.985$), and chamber-inferred networks - consistent with these biological findings	[32]
LIANA+ , CellChat, CellPhoneDB	Benchmarked frameworks; uses multi-method ligand- receptor inference, NMF decomposition, and variance explained (R^2), identifies FB-to- myeloid and cardiac signalling axes	HeartMAP integrates LIANA for ligand- receptor inference; reports chamber- specific and myeloid/ fibroblast interaction hubs similar to LIANA+ multi- method outputs	[21]
CellChat	Inferred quantitative cell-cell communication network; measured centrality, pathway enrichment, dominant pairs in heart tissue	HeartMAP offers communication networks, pathway enrichment, and hub scores; outputs overlap in structure and biological relevance with CellChat	[5]
Ligand–receptor interaction	Identified 120 unique ligand-receptor pairs in nine heart cell types (single-cell RNA-seq)	HeartMAP identifies basic communication networks and differentially expressed genes per chamber, behaving similarly to ligand- receptor mapping approaches	[33]
CellPhoneDB, CellChat	Cell–cell interactions via ligand–receptor networks between cardiomyocytes and non-myocytes; spatial transcriptomics for spatial context	HeartMAP incorporates spatial transcriptomics and detects similar cell- cell communication structures and context-aware signalling axes	[34]

Thanyani Pandelani: Supervision.

Data and Code Availability

All scripts, results, and documentation are available in the HeartMAP repository at <https://github.com/Tumo505/HeartMap>. The original dataset is accessible through the Single Cell Portal (SCP498) at https://singlecell.broadinstitute.org/single_cell/study/SCP498/transcriptional-and-cellular-diversity-of-the-human-heart#study-summary. Processed chamber-specific results, communication networks, and visualisation tools are made available for community use and further research. The file structure includes analysis scripts for basic pipeline, advanced communication, and multi-chamber atlas analyses, along with detailed results and publication-ready figures.

Consent for Publication

Not applicable. This study used publicly available, de-identified datasets where consent for publication was obtained by the original data collectors.

Guarantor

TP

Contributorship

TK conceived the HeartMAP framework, designed the methodology, performed computational analysis and drafted the manuscript. LW contributed to computational analysis methods and critical review of computational and statistical approaches. HN supported results interpretation, statistical and methodological refinement. TP provided overall project supervision and ensured scientific rigor throughout the research process. All read and approved the final manuscript.

Author Agreement Statement

We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us. We understand that the Corresponding Author is the sole contact for the Editorial process. He is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs

Declaration of Competing Interest

The authors declare that there are no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgements

We acknowledge the Single Cell Portal (SCP498) for providing the human heart single-cell RNA-sequencing dataset that enabled this multi-chamber analysis.

References

[1] J.A. Palmer, N. Rosenthal, S.A. Teichmann, and M. Litvinukova, “Revisiting Cardiac Biology in the Era of Single Cell and Spatial Omics,” Jun. 2024. doi: 10.1161/CIRCRESAHA.124.323672.

[2] Chen Y, Liu Y, Gao X. The Application of Single-Cell Technologies in Cardiovascular Research. Front Media S A Oct. 2021. <https://doi.org/10.3389/fcell.2021.751371>.

[3] Yamada S, Nomura S. Review of single-cell RNA sequencing in the heart. MDPI AG Nov. 2020. <https://doi.org/10.3390/ijms21218345>.

[4] Wang S, Zheng H, Choi JS, Lee JK, Li X, Hu H. A systematic evaluation of the computational tools for ligand-receptor-based cell–cell interaction inference. Oxford University Press; Sep. 2022. <https://doi.org/10.1093/bfpg/ela019>.

[5] Jin S, et al. Inference and analysis of cell-cell communication using CellChat. Nat Commun Dec. 2021;12. <https://doi.org/10.1038/s41467-021-21246-9>.

[6] Martins-Marques T. Connecting different heart diseases through intercellular communication. Co Biol Ltd Sep. 2021. <https://doi.org/10.1242/bio.058777>.

[7] Fountoulaki K, Dagres N, Iliodromitis EK. Cellular Communications in the Heart. Card Fail Rev Oct. 2015;1:64–8. <https://doi.org/10.15420/cfr.2015.1.2.64>.

[8] Iacobas S, Amuzescu B, Iacobas DA. Transcriptomic uniqueness and commonality of the ion channels and transporters in the four heart chambers. Sci Rep Dec. 2021; 11. <https://doi.org/10.1038/s41598-021-82383-1>.

[9] S.Y. Ng, C.K. Wong, and S.Y. Tsang, “Differential gene expressions in atrial and ventricular myocytes: Insights into the road of applying embryonic stem cell-derived cardiomyocytes for future therapies,” Dec. 2010. doi: 10.1152/ajpcell.00402.2009.

[10] Tucker NR, et al. Transcriptional and Cellular Diversity of the Human Heart. Circulation Aug. 2020;142:466–82. <https://doi.org/10.1161/CIRCULATIONAHA.119.045401>.

[11] Ackers-Johnson M, Tan WLW, Foo RSY. Following hearts, one cell at a time: recent applications of single-cell RNA sequencing to the understanding of heart disease. Nat Publ Group Dec. 2018. <https://doi.org/10.1038/s41467-018-06894-8>.

[12] Martins-Marques T. Connecting different heart diseases through intercellular communication. Co Biol Ltd Sep. 2021. <https://doi.org/10.1242/bio.058777>.

[13] Zhao XS, Gallardo TD, Lin L, Schageman JJ, Shohet RV. Transcriptional mapping and genomic analysis of the cardiac atria and ventricles. Physiol Genom 2003;12: 53–60. <https://doi.org/10.1152/physiolgenomics.00086.2002>.

- [14] R.C. Scott, D. Crabbe, B. Krynska, R. Ansari, and M.F. Kiani, "Aiming for the heart: Targeted delivery of drugs to diseased cardiac tissue," Apr. 2008. doi: [10.1517/17425247.5.4.459](https://doi.org/10.1517/17425247.5.4.459).
- [15] Lickiss B, et al. Chamber-specific contractile responses of atrial and ventricular hiPSC-cardiomyocytes to GPCR and ion channel targeting compounds: A microphysiological system for cardiac drug development. J Pharm Toxicol Methods Jul. 2024;128. <https://doi.org/10.1016/j.vascn.2024.107529>.
- [16] Zhang Z, Villalpando J, Zhang W, Nam YJ. Chamber-specific protein expression during direct cardiac reprogramming. Cells Jun. 2021;10. <https://doi.org/10.3390/cells10061513>.
- [17] Tucker NR, et al. Transcriptional and Cellular Diversity of the Human Heart. Circulation Aug. 2020;142(5):466. <https://doi.org/10.1161/CIRCULATIONAHA.119.045401>.
- [18] "Transcriptional and Cellular Diversity of the Human Heart - Single Cell Portal." Accessed: Jul. 23, 2025. [Online]. Available: (https://singlecell.broadinstitute.org/single_cell/study/SCP498/transcriptional-and-cellular-diversity-of-the-human-heart#study-summary).
- [19] He J, Lin L, Chen J. Practical bioinformatics pipelines for single-cell RNA-seq data analysis. Biophys Rep 2022;8(3):158. <https://doi.org/10.52601/BPR.2022.210041>.
- [20] Luecken MD, Theis FJ. Current best practices in single-cell RNA-seq analysis: a tutorial. Mol Syst Biol Jun. 2019;15(6):8746. https://doi.org/10.15252/MSB.20188746/SUPPL_FILE/MSB188746-SUP-0003-DATASET1.V1.ZIP.
- [21] Dimitrov D, et al. LIANA+ provides an all-in-one framework for cell-cell communication inference. Nat Cell Biol Sep. 2024;26(9):1613. <https://doi.org/10.1038/S41556-024-01469-W>.
- [22] S. Jin, M.V. Plikus, and Q. Nie, "CellChat for systematic analysis of cell-cell communication from single-cell transcriptomics," *Nat Protoc*, vol. 20, no. 1, pp. 180–219, Jan. 2025, doi: [10.1038/S41596-024-01045-4](https://doi.org/10.1038/S41596-024-01045-4); SUBJMETA=114,2391,2694,553,631,794;KWRD=CELLULAR+SIGNALLING+NETWORKS,MULTICELLULAR+SYSTEMS,SOFTWARE.
- [23] Wu H, Yang Z, Cao Q, Wang P, Biswal BB, Klugah-Brown B. MQGA: A quantitative analysis of brain network hubs using multi-graph theoretical indices. Neuroimage Dec. 2024;303. <https://doi.org/10.1016/j.neuroimage.2024.120913>.
- [24] Salavaty A, Ramialison M, Currie PD. Integrated Value of Influence: An Integrative Method for the Identification of the Most Influential Nodes within Networks. Patterns Aug. 2020;1(5):100052. <https://doi.org/10.1016/J.PATTERN.2020.100052>.
- [25] Kane C, Terracciano CMN. Concise Review: Criteria for Chamber-Specific Categorization of Human Cardiac Myocytes Derived from Pluripotent Stem Cells. Stem Cells Aug. 2017;35(8):1881. <https://doi.org/10.1002/STEM.2649>.
- [26] Mou T, Deng W, Gu F, Pawitan Y, Vu TN. Reproducibility of Methods to Detect Differentially Expressed Genes from Single-Cell RNA Sequencing. Front Genet Jan. 2020;10:1331. <https://doi.org/10.3389/FGENE.2019.01331/FULL>.
- [27] Liang S, Liang Q, Chen R, Chen K. Stratified Test Alleviates Batch Effects in Single-Cell Data. Algorithms Comput Biol 2020;12099:167. https://doi.org/10.1007/978-3-030-42266-0_13.
- [28] Haynes W. Benjamini-Hochberg Method," in. Encyclopedia of Systems Biology. Springer New York; 2013. p. 78. https://doi.org/10.1007/978-1-4419-9863-7_1215.
- [29] Shao X, Lu X, Liao J, Chen H, Fan X. New avenues for systematically inferring cell-cell communication: through single-cell transcriptomics data. Protein Cell Dec. 2020;11(12):866. <https://doi.org/10.1007/S13238-020-00727-5>.
- [30] Benesty J, Chen J, Huang Y, Cohen I. Pearson correlation coefficient. Noise reduction in speech processing. Springer; 2009. p. 1–4.
- [31] Rao M, et al. Resolving the intertwining of inflammation and fibrosis in human heart failure at single-cell level. Basic Res Cardiol Dec. 2021;116(1). <https://doi.org/10.1007/S00395-021-00897-1>.
- [32] Litvinuková M, et al. Cells of the adult human heart. Nat 2020 5887838 Sep. 2020; 588(7838):466–72. <https://doi.org/10.1038/s41586-020-2797-4>.
- [33] Zhang L, et al. Identification of cell-to-cell interactions by ligand-receptor pairs in human fetal heart. Biochim Et Biophys Acta (BBA) Mol Basis Dis Dec. 2020;1866(12):165917. <https://doi.org/10.1016/J.BBADS.2020.165917>.
- [34] Tan S, Yang J, Hu S, Lei W. Cell-cell interactions in the heart: advanced cardiac models and omics technologies. Stem Cell Res Ther Dec. 2024;15(1):362. <https://doi.org/10.1186/S13287-024-03982-Z>.

HeartMAP: A multi-chamber spatial framework for cardiac cell-cell communication

Tumo Kgabeng, Lulu Wang, Harry Ngwangwa, and Thanyani Pandelani

Citation: Kgabeng T, Wang L, Ngwangwa H, Pandelani T. HeartMAP: A multi-chamber spatial framework for cardiac cell-cell communication. *Comput Struct Biotechnol J*. 2025;27:4976-4991. DOI: 10.1016/j.csbj.2025.11.015

Abstract

Understanding cell-cell communication within and between the four distinct cardiac chambers is fundamental to elucidating cardiac function and disease mechanisms. Each chamber exhibits unique cellular and molecular characteristics that reflect specialised physiological roles, yet existing frameworks for mapping chamber-specific intercellular networks have remained limited. Here, we present HeartMAP (Heart Multi-chamber Analysis Platform), a computational framework that infers cardiac cell-cell communication networks at chamber resolution through integration of single-cell RNA-seq co-expression patterns and ligand-receptor interaction databases. Using a dataset of 287,269 cells from seven healthy human heart donors (Single Cell Portal SCP498), we identified chamber-specific cell populations, communication networks, and therapeutic targets. HeartMAP employs a progressive three-tier analytical approach comprising basic pipeline analysis, advanced communication modelling and multi-chamber atlas construction to reveal both conserved and chamber-specific signalling pathways; cross-chamber correlation analysis demonstrated the highest similarity between ventricles ($r=0.985$) and the lowest between left atrium and left ventricle ($r=0.870$), reflecting functional specialisation. Communication hub analysis identified atrial cardiomyocytes and adipocytes as key signalling centres with hub scores of 0.037–0.047, while differential expression analysis revealed over 150 significantly different genes per chamber pair. These findings establish a molecular foundation for precision cardiology approaches, enabling chamber-specific therapeutic strategies that could improve treatment outcomes for cardiovascular diseases. HeartMAP is freely available as a Python package than can be installed using “pip install heartmap”, the package’s documentation can be found on <https://pypi.org/project/heartmap/>, it can also be accessed via a user-friendly web interface freely available at <https://huggingface.co/spaces/Tumo505/heartmap-cell-analysis>.

Graphical Abstract

View the article online

<https://spj.science.org/doi/10.1016/j.csbj.2025.11.015>

Use of this article is subject to the [Terms of service](#)

(ISSN) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005.

© 2025 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).