Denver Ncube, Ph. D

Morrisville, NC • +1 541-606-7565

denver.ncube@gmail.com https//linkedin.com/in/denver-ncube-n3d/ https://github.com/DenverN3

TECHNICAL EXPERTISE

- NGS data analysis with R, Python & Linux (scRNAseq and bulk RNAseq, WGS, WES)
- Whole Genome and Whole Exome Seq analysis (GATK, vcf, bcf tools, picard, Samtools)
- Spatial transcriptomic data analysis (Nanostring GeoMx and 10X Visium, MERFISH, Vizgen, Seurat, Monocle, scanPy, squidPy,RNA Velocity, scVelo)
- Multi-Omics data integration, metabolomics data analysis
- Bioinformatics pipeline development and deployment (CI-CD)
- Data analysis with R, R-Shiny, Python(Jupyter Notebook,pandas, numpy,matplotlip, seaborn), Unix, & SPSS
- Cloud computing (AWS S3 and EC2)
- Artificial Intelligence integration in Bioinformatics workflows
- Machine Learning with Python(scikit-learn, rdkit) & R (Supervised and Unsupervised)
- NextFlow and Snakemake pipeline construction
- RWE data analysis in oncology and IBD

TRANSFERABLE SKILLS

- Innovation
- Cross-functional team leadership & collaboration
- Adaptable and flexible learner
- Leadership
- Training and Mentorship
- Oncology and Inflammatory Bowel Disease TA expertise
- Project Management
- Client Liaison
- Business strategy
- Precision medicine investment advisory
- Good clinical practice (GCP) and GDPR certification
- NIH information management awareness certification
- RUO and CLIA level assay validation expertise
- CAP-CLIA regulatory experience

Education

University of Oregon

Ph.D.: Biology(Neuroscience) Eugene, OR • 2021

WORK EXPERIENCE

Precision Medicine Consultant (Part-time) — Theriome, 02-2025 - present

- Review and interpret customer Aristotle Test reports with precision and clarity.
- Conduct insightful 30-minute Zoom consultations, providing personalized health insights based on test results.
- Continuously enhancing the company's expertise in systems biology and bioinformatics through implementing high end data analytics and AI-integrated solutions.

Clinical Translational Scientist — Alimentiv, 06-2024 - 02-2025

- Provided expert input on design and conduct of clinical studies incorporating Precision Medicine endpoints and or omics and non-omics techniques including but not limited to transcriptomics (bulk, single cell and spatial-10X Visium), metabolomics, proteomics, microbiome, digital image analysis and or clinical pharmacology.
- Utilizing multi-omics integration tools (Mulit-Omics Factor Analysis) to integrate omics and non-omics data
- Integration of RWD-RWE such as EHR into the transcriptomic analysis (bulk scRNAseq and Spatial Transcriptomics) of data from phase 1-2a clinical trials in Inflammatory Bowel Disease, Chron's Disease and Ulcerative Colitis.
- Created comprehensive technical documentation and user manuals for internal pipelines for R facilitating cross-team reproducibility.
- Designed translational Statistical Analysis Plans for bulk and spatial transcriptomics analysis of clinical trial data.
- Prepared and supported the preparation and review of Precision Medicine product plans, clinical study reports, and Precision Medicine sections of clinical development plans, study protocols, abstracts, manuscripts, and technical reports.
- Performed and supported the analysis and interpretation of Precision Medicine data for sponsor and investigator initiated studies with statistical analysis tools like R Studio.
- Standardized version-controlled code repositories using GitHub ensuring reproducilble workflows and collaboration with the Precision Medicine Team
- Prepared analytical product improvement plans to enhance client offerings and optimizing data processing.
- Developed an internal repository through integrating existing publication collections and utilizing LLM based solutions on AWS like Amazon Q for Business.
- Performed Software Life-cycle Testing and validation of analytical infrastructure to enhance high throughput analysis of multiomic data.
- Facilitated the migration of data analysis pipelines from on-premise platforms to AWS.
- Assisted with business development activities which included participating in Request for Proposal (RFP) responses, project pricing, protocol development and bid defenses in the area of Precision Medicine.
- Provided advice to sponsors seeking Alimentiv's guidance in early clinical phase work.
- Ensured traceability and reproducibility of analyses in support of clinical study reports and potential regulatory submissions
- Assisted in the development of product and service platforms to ensure all stakeholder requirements were being considered, documented, and translated into the final solution.

Precision Medicine Consultant (Part-time) — **Theriome**, 02-2024 - 07-2024

- Review and interpret customer Aristotle Test reports with precision and clarity.
- Preparation of consultation materials tailored to each customer's unique profile.
- Conduct insightful 30-minute Zoom consultations, providing personalized health insights based on test results.
- Continuously enhancing the company's expertise in systems biology and bioinformatics.
- Participate in customer acquisition through creative networking and marketing strategies tailored for KOLs in the precision medicine field.
- Developed a CLIA validation protocol for the Aristotle test, charting the way towards full validation of the test.

Biological Data Expert (Part-time - as needed) **Mercor Intelligence,** 07-2024 -present

- Training and reviewing performance of LLM models on advanced biology and genetics content.
- Evaluating LLM performance on complex reasoning and mathematical tasks
- Setting benchmarks for LLM performance through rigorous monitoring of model execution of tasks by comparing with competitor LLMs.

Human AI Data Trainer (Part-time) **OpenAI,** 02-2024 - 07- 2024

- Training and reviewing performance of LLM models on advanced human biology and genetics

Precision Medicine Consultant (Part-time) **GLG,** 02-2024 - 07- 2024

- Conducting consults on a broad range of strategic issues of Artificial Intelligence in the Biotech and Biopharmaceutical sectors for drug discovery and multi-modal data.
- Providing strategic input on precision medicine investment and cutting edge market research data and important regulatory frameworks governing the same.

Computational Biologist I **Flagship Biosciences Morrisville, NC, USA** • 03-2023 - 01-2024

- Developed, updated and oversaw additional RNAseq workflows related to 5 new NGS product initiatives.
- Developed **c**ustom pipelines for and for next generation sequencing data analysis (spatial transcriptomics data normalization and visualization) of oncology and other data which adhering to CLIA/CAP guidelines.
- Developed clinical informatics pipelines for somatic variant curation in myeloid cancer for companion diagnositc panel validation.
- Conducted genetic variant analysis including SNP/indel calling from WGS and RNAseq data with GATK, bcftools, Samtools,snpEFF etc to explore disease-related gene variation.
- Analyzed and curated variants for companion diagnostic panels cross referencing with known human databases such as dbSNP, ClinVar, OMIM, gNomad in line with ACMG classification criteria.
- Applied population stratification and genetic association analysis techniques to support biomarker correlation studies across different clinical cohorts.
- Collaborated with clinical data managers and regulatoy teams to prepare bioinformatics documentation in alignment with GxP and FDA-compliant data analysis standards.
- Utilized advanced statistical analysis for multi-modal (EHR and clinical data) data analysis to enhance clinical insight.
- Analyzed RNAseq (single cell RNAseq, bulk RNASeq) and DNAseq (WES, WGS, Amplicon) data using Illumina BaseSpace Sequencing Hub tools such as DRAGEN RNA, DRAGEN DNA Amplicon for RUO and CLIA purposes.
- Analyzed multi-omic data utilizing network biology databases such as PPI, GSEA, Swiss-Prot, EMBL, KEGG etc.
- Assisting in the development and expansion of Flagship's "data as a service" platform through integrated data analysis pipeline development.
- Preparing and integrating different types of omics data (QC, Batch correction) including data from oncology bulk, single cell RNAseq, Whole Genome(WGS) & Whole Exome Sequencing(WES), and spatial transcriptomics experiments and imaging data to enhance clinical insight.
- Leading custom oncology companion diagnostics screening development work in solid tumors through intelligent data mining of publicly available human genome repositories such as TCGA, GEO, dbSNP, OMIM, ClinVar, cBioPortal, ENCODE, Ensembl, NCBI, and UCSC databases.
- Leading data analysis and visualization tool development (R-Shiny app design) for all new NGS based assays in development currently offered to our various clients.
- Deploying algorithms and statistical methods to extract meaningful insight from spatial multi-omics data from NanoString GeoMx platform.
- Led analysis for 4 RUO and CLIA assay validation for custom oncology panels with R, Python, DRAGEN, and Alissa Interpret using both synthetic and patient (RWD) samples.
- Implemented project tracking through HIVE project tracking software streamlining internal and external bioinformatics requests increasing turnaround time by 80%.
- Instituted a report generation protocol for all in-house bioinformatics requests and storage to show the business contribution of the Computational Sciences Department.
- Optimized raw NGS data processing through the use of customized python pipelines reducing client data delivery time by at least 60%.
- Instituted a responsive bioinformatics tools development program as part of the Assay Development Team.

Data Scientist, Self-Employed(Pro-Bono-Unpaid) **Mount Holly, NC** 08-2021 - 03-2023

- Led the project management team for implementation of Natural Language Processing (NLP) for Dialect Dialogue African language preservation project.
- Assisted in the writing of grant proposals for integration of VR or AI for grade level and high school students.
- Designed a data-based app for optimization of bid submission by Wise Cargo and Blackrock Expedited Shipping services increasing profitability by 45%.
- Providing bio-informatics technical support for researchers and graduate students.
- Analyzed bulk and single cell RNAseq data from human and zebrafish samples using both R and Python software packages.
- Analyzed data from chemo-informatic and bioinformatic databases such as CHEMBL, UCSC Genome browser NCBI, and Ensembl for various clients.
- Analyzed Whole Genome Sequencing data and performed variant calling with Unix based tools, FastQC, SRA tools, Samtools and GATK to identify SNPs and InDels.
- Optimized RNAseq data analysis pipelines adhering to best practices using common workflow languages (CWL) such as Snakemake and NextFlow.

Research Scientist, University of Oregon Institute of Neuroscience Eugene OR, 06-2021 - 08-2021

- Analyzed and curated bulk and single cell RNAseq data from zebrafish multi-transmitter neurons using R-Studio and Python bio-informatics packages (BioConductor, STAR, DESEq2, TopGO,Seurat, GATK, picard, samtools etc) for 3 publications.
- Utilized various data visualization packages in R (R-Shiny, Plotly, ggplot) and Python(Seaborn) to enhance extraction of insights from bulk and single cell sequencing.
- Discovered and published on a “transcriptomic fingerprint” through analysis of neuronal RNAseq data to map a cluster of neurons from the zebrafish to mouse brain.
- Processed raw NGS data in Unix environment in preparation for downstream analysis for over 20 experiments.
- Designed and carried out transcriptional profiling of multi-transmitter neurons after FACSorting and sequencing on the Illumina Hi-Seq Platform.
- Analyzed and processed raw sequencing data to generate count matrices downstream differential expression analysis.
- Investigated the impact of gut microbiota on multi-transmitter neuron specification and gene expression using single cell and bulk-RNAseq data.
- Optimized fluorescent in-situ hybridization(FISH), immuno-histochemistry(IHC),confocal microscopy imaging protocols increasing experimental efficiency.
- Prepared and published manuscripts on transcriptional profiling of multi- transmitter neurons in the zebrafish forebrain.
- Optimized record keeping through eLN systems such as Labarchive.
- Managed complex resources including paperwork, ordering, protocols, SOPs, experiment materials.
- Prepared, images, tables and graphs to summarize research results for publication.

Graduate Research Scientist, University of Oregon Institute of Neuroscience Eugene OR, 09-2016 - 06-2021

- Analyzed and curated bulk and single cell RNAseq data from zebrafish multi-transmitter neurons using R-Studio and Python bio-informatics packages (BioConductor, STAR, DESEq2, TopGO,Seurat, GATK, picard, samtools etc) for research publications and thesis.
- Utilized bioinformatic platforms (Benchling, Geneious,Ensembl(BLAT,BLAST etc), Genome Browser, CCTOP, CRISPRScan) in optimizing the design of single guide RNA.
- Optimized multi-plexed CRISPR-Cas9 gene editing to knockout genes in reverse genetic screens in zebrafish.
- Deployed phenotype screening pipelines for assessing CRISPR off-targets using RNAseq data analysis with common Linux or Unix based NGS data analysis tools.
- Delivered gene expression constructs in mammalian cells (COS7 and MDCK) to investigate chemical synapse formation.
- Designed in-vitro and in-vivo assays to determine efficacy of CRISPR-Cas9 gene editing.
- Utilized knowledge of CRISPR based genome editing technologies such as, deletions, knock-ins to optimize sgRNA design.
- Investigated chemical synapse formation using standard nanoparticle transfection of mammalian cells with plasmids carrying cloned genetic sequences as well as CRISPR-Cas9 gene editing.
- Discovered and published on a“transcriptomic fingerprint” through analysis of neuronal bulk and single cell RNAseq data to map a cluster of neurons from the zebrafish to mouse brain.
- Designed data analysis pipelines for single cell and bulk RNA sequencing data including data visualization tools such as dashboards with R-Shiny, ggplot and plotly.
- Extracted, isolated and purified nucleic acids for sgRNA and FISH probe synthesis reducing task time by 50%.