**Required Packages**

**MATLAB 9.10+**

**Bioinformatics Toolbox 4.15.1**

**Symbolic Math Toolbox 8.7**

**(8.77GB)**

**Python 3.9+**

**Main TrueProbe Probe Design**

* **A0 Probe Generation** 
  + **MATLAB 9.10**
  + **Bioinformatics Toolbox 4.15.1**
  + **MATLAB Parallel Server 7.4**
  + **Polyspace Bug Finder 3.4**
  + **Polyspace Bug Finder Server 3.4**
* **A1 Probe Blasting**
  + **MATLAB 9.10**
  + **Bioinformatics Toolbox 4.15.1**
  + **Parallel Computing Toolbox 7.4**
  + **MATLAB Parallel Server 7.4**
  + **Polyspace Bug Finder 3.4**
* **A2 GeneExpressionInformation**
  + **MATLAB 9.10**
  + **Statistics and Machine Learning Toolbox 12.1**
  + **Parallel Computing Toolbox 7.4**
  + **MATLAB Parallel Server 7.4**
  + **Polyspace Bug Finder 3.4**
* **A3 ThermodynamicInformation**
  + **MATLAB 9.10**
  + **Symbolic Math Toolbox 8.7**
  + **Statistics and Machine Learning Toolbox 12.1**
  + **Bioinformatics Toolbox 4.15.1**
  + **Parallel Computing Toolbox 7.4**
  + **MATLAB Parallel Server 7.4**
  + **Polyspace Bug Finder 3.4**
* **A4 BindingSiteMapping**
  + **MATLAB 9.10**
  + **Statistics and Machine Learning Toolbox 12.1**
  + **Bioinformatics Toolbox 4.15.1**
  + **Parallel Computing Toolbox 7.4**
  + **MATLAB Parallel Server 7.4**
  + **Polyspace Bug Finder 3.4**
* **A5 ProbeDesignerStats**
  + **MATLAB 9.10**
* **A6 ProbeSelection**
  + **MATLAB 9.10**
  + **Bioinformatics Toolbox 4.15.1**
  + **Parallel Computing Toolbox 7.4**
  + **MATLAB Parallel Server 7.4**
  + **Polyspace Bug Finder 3.4**
* **A7 GetMetrics**
  + **MATLAB 9.10**
  + **Statistics and Machine Learning Toolbox 12.1**
  + **Bioinformatics Toolbox 4.15.1**
* **A8 ProbeSpecificityFiltering**
  + **MATLAB 9.10**
  + **Bioinformatics Toolbox 4.15.1**
* **A9 MultiFluorphoreColocalizationConvolutionND**
  + **MATLAB 9.10**
  + **Symbolic Math Toolbox 8.7**
  + **Image Processing Toolbox 11.3**

**Run Other Probe Design**

* **GenerateIndividualGeneFasta**
  + **MATLAB 9.10**
  + **Bioinformatics Toolbox 4.15.1**
* **Run OtherProbeSets GetMetrics**
  + **MATLAB 9.10**
  + **Statistics and Machine Learning Toolbox 12.1**
  + **Bioinformatics Toolbox 4.15.1**

**Compare Software Probe Design**

* **IntegrateResults**
  + **MATLAB 9.10**
  + **Bioinformatics Toolbox 4.15.1**
* **PrintProbeSets**
  + **MATLAB 9.10**
  + **Bioinformatics Toolbox 4.15.1**
* **CalculateTemperatureOrNewConditionSolution**
  + **MATLAB 9.10**
* **Main PlotResults**
  + **MATLAB 9.10**
  + **Symbolic Math Toolbox 8.7**
  + **Image Processing Toolbox 11.3**
  + **Statistics and Machine Learning Toolbox 12.1**
  + **Curve Fitting Toolbox 3.5.13**
  + **Parallel Computing Toolbox 7.4**
  + **MATLAB Parallel Server 7.4**
  + **Polyspace Bug Finder 3.4**

Running software.

Software Has Main Script.

A0\_BKJH\_ProbeDesign\_Wrapper\_cluster\_V5(id,cluster)

For it need to give main input for genes designed.

inputs1 = {...

{'NM\_000805.5'},{},{}, 'Human','(GAST)','(GAST)','17',{},1 ;...

For this input Accession Number, Organism, Gene Name, Chromosome, and if want sequence or anti-sense sequence.

**Input Parameters.**

* **Locations**
  + **SaveRoot**
    - [Location where files are to be saved]
  + **customBlastDatabase\_DNA**
    - [Location of user-custom DNA blast database if user wants to use their own custom database]
  + **customBlastDatabase\_RNA**
    - [Location of user-custom RNA blast database if user wants to use their own custom database]
  + Loc
  + LocRoot
  + Human\_wgEncodeGencodeRefSeqFile
  + Human\_wgEncodeGencodeAttributesFile
  + Human\_wgEncodeGencodeCompFile
  + Human\_GencodeRefSeqMetadataFile
  + Human\_GenomeAssemblyReportFile
* **SingleOverMultiplex** 
  + - [if design is for a single gene (1) or multiplexed genes (0), default 1]
* **AllIsoforms** 
  + - [if design is all gene isoforms (1) or designed the specific isoform specified (0), default 0]
* **Main Probe constraints**
  + **max\_probes**
    - [Max number of probes to design, default 96]
  + **minProbeSize**
    - [min nt length of potential probes, default 20]
  + **maxProbeSize**
    - [max nt length of potential probes, default 20]
  + **MininumProbeSpacing**
    - [min spacing between probes, default 3]
* **Thermodynamic constraints**
  + **HybridizationTemperature** 
    - [Hybridization temperature, default 37C]
  + **SaltConcentration**
    - [Salt Concentration mM, default 0.05]
* **BLAST Parameters**
  + **BLASTrna**
    - [decide to blast RNA sequences, default 1]
  + **BLASTdna**
    - [decide to blast DNA sequences, default 0]
* **Parallelization Parameters**
  + **batchSize**
    - [number of probes to blast/evaluate at a time, default 10]
  + **targetSize**
    - [number of targets to compute at a time, default 200]
  + **ParsingPreference** 
    - [blast simultaneously in parallel(1) or blast probes sequentially (0), default 1]
* **Expression Parameters**
  + **DoAllGenesHaveSameExpression**
    - [decide to assume equal expression for all genes (1) or to use gene expression reference (0) , default 1]
  + **nullRNAcopynumber**
    - [number of RNA copy when not using reference expression levels, default 100]
  + **nullDNAcopynumber**
    - [number of DNA copy number when not using reference expression levels, default 2]
  + **UseGeneOverTranscLevelExpression** 
    - [use gene level (1) or transcript isoform level (0) gene expression values, default 0]
  + **MouseExpressionFile**
    - [location of mouse gene expression bed file]
  + **HumanTCGA\_TranscriptExpressionFile**
    - [location of human TCGA transcript expression bed file]
  + **HumanTCGA\_GeneExpressionFile**
    - [location of human TCGA gene level expression bed file]
  + **HumanGTEX\_TranscriptExpressionFile**
    - [location of human GTEx transcript expression bed file]
  + **HumanGTEX\_GeneExpressionFile**
    - [location of human GTEx gene level expression bed file]
* **RNA Secondary Structure Parameters**
  + **SolveSecondaryStructure** 
    - [decides to compute RNA secondary structure denovo, default 0]
  + **SecondaryStructureFileRoot** 
    - [Location of preknown secondary structures files]

**Outputs Files. NLPDS (Neuert Lab Probe Design Software).**

* **(GeneName)\_RefSeqID\_probes\_NLPDS.mat**
  + [structure with probe sequences, location on on-target]
* **(GeneName)\_RefSeqID \_hits\_table\_NLPDS.mat**
  + [structure with information on BLAST hits]
* **(GeneName)\_RefSeqID\_ExpressionInfo\_NLPDS.mat**
  + [Structure with expression in TCGA and GTEX]
* **(GeneName)\_RefSeqID\_Tm37\_OnOffThermoInfo\_NLPDS.mat**
  + [Structure with binding energy of all hits]
* **(GeneName) \_RefSeqID \_dCpInfo\_NLPDS.mat**
  + [Structure with heat capacity for all target binding reactions]
* **(GeneName)\_RefSeqID\_dHInfo\_NLPDS.mat**
  + [Structure with enthalpy for all target bindings reactions]
* **(GeneName)\_RefSeqID\_dSInfo\_NPLDS.mat**
  + [Structure with entropy for all target binding reactions]
* **(GeneName)\_binding\_hits\_map\_NPLDS**
  + [binding site map]
* **(GeneName)\_RefSeqID\_Tm37\_BindingEnergyMatrix\_NPLDS.mat**
  + [Equilibrium Binding Energy in binding site map format]
* **(GeneName)\_RefSeqID\_BindingMatricies\_NPLDS.mat**
  + [Entropy, Enthalphy, and heat capacity in binding site map format for RNA]
* **(GeneName)\_RefSeqID\_BindingMatricies\_NPLDS.mat**
  + [Entropy, Enthalphy, and heat capacity in binding site map format for complementary strand DNA binding]
* **(GeneName)\_RefSeqID\_Tm37\_BasicDesignerStats\_NPLDS.mat**
  + [Index information on stats used for design probes]
* **(GeneName)\_RefSeqID\_chosen.mat**
  + [List of chosen probe indexes[
* **(GeneName)\_RefSeqID\_probes\_final\_96max.xlsx**
  + [Excel spreadsheet with final probes, and some stats]
* **(GeneName)\_RefSeqID\_Tm37\_ModelMEtrics\_NLPDS.mat**
  + [Structure with binding affinity calculations and probe design metrics]

gene\_num = id;

cellPreset = 1;

updateLocation = 1;

currLoc = 2+cluster;

RemoveMisMatches = 1;

SpecificityThreshold = 2;

DecisionAlgorithmFloorSize = 0.5;

RemoveRibosomalHits = 1;

RunOffline = 1;

withNascentTranscripts = 0;

settings.clusStat = cluster;

A0\_BKJH\_ProbeDesign\_Wrapper\_cluster\_V5(id,cluster)

Main Options:

Design for Single Isoform, or across multiple isoforms: (Isoform-Resolved, Isoform-Flattened)

MATLAB 2023b

First Install Python 3.9 and add python.exe to PATH?