As of September 22, 2020

Overview of LRGASP Challenges

Challenge 1: Transcript isoform detection with a high-quality genome
Goal: Identify which sequencing platform, library prep, and computational tool(s) combination gives the highest sensitivity and precision for transcript detection.

Challenge 2: Transcript isoforms quantification
Goal: Identify which sequencing platform, library prep, and computational tool(s) combination gives the most accurate expression and expression fold-change estimates.

Challenge 3: De-novo transcript isoform detection without a high-quality genome
Goal: Identify which sequencing platform, library prep, and computational tool(s) combination gives the highest sensitivity and precision for transcript detection.

Evaluation of submissions will follow the procedure established by the LRGASP consortium. There are many metrics for evaluation and one tool or one method may not perform best at all metrics.

Challenge 1 Evaluation: Transcript isoform detection

Evaluation of transcriptome annotation for Human and Mouse models

Four sets of transcripts will be used for evaluation of transcript calls

1. Lexogen SIRV Set 4 (SIRV Set 3 plus 15 new long SIRVs with sizes ranging from 4KB-12KB)
2. Comprehensive GENCODE annotation, latest version when LRGASP submissions are evaluated. May be newly released after submissions (Human v36, Mouse vM26, if available. If not Human v35, Mouse vM25. Human genome assembly: GRCh38. Mouse genome assembly: GRCm39)
3. A subset of undisclosed, manually curated transcripts by GENCODE considered as “bona fide” derived from LRGASP data
4. Simulated data (Trans-Nanosim, Iso-SeqSim)

The evaluation will use SQANTI and categories that will serve as a basis to compute LRGASP evaluation metrics detailed below. The evaluation script will be provided.

SQANTI Transcript Classifications

Classification	Description
Full Splice Match (FSM)	Transcripts matching a reference transcript at all splice junctions
Incomplete Splice Match (ISM)	Transcripts matching consecutive, but not all, splice junctions of the reference transcripts
Novel in Catalog (NIC)	Transcripts containing new combinations of already annotated splice junctions or novel

	splice junctions formed from already annotated donors and acceptors.
Novel Not in Catalog (NNC)	Transcripts using novel donors and/or acceptors

A number of novel transcripts detected by all or most pipelines, as well as pipeline-specific transcripts will be selected for experimental validation and manual review by the GENCODE project.

A pilot to demonstrate evaluation metrics was performed and summary slides can be found here.

Evaluation by *SIRVs*

TP: Number of FSM with 3' and 5' ends within 50 nts of annotated TSS and TTS, respectively, with transcript models FULLY supported by full-length reads. Only one TP per SIRV model counts.
FN: Number of SIRVs - TP
FP: Number of ISM, NIC, NNC transcripts matching a SIRV annotation
Sensitivity: TP/ number of SIRVS
Precision: TP/ transcripts mapped to a SIRV transcript
FDR: FP/ transcripts mapped to a SIRV transcript

Evaluation by Comprehensive GENCODE annotation

FSM transcripts (the transcript matches all junctions of a reference transcript)
True Positives (TP): FSM with TSS and TTS within 50nts distance from their reference match
3' end True (3'T): FSM with polyA site
5' end True (5'T): FSM with CAGE support
All True Positives (AllTP): FSM with 5' end within 50nts distance from reference TSS or CAGE support AND 3' end within 50nts distance from reference TTS or polyA site prediction.
Normalized True Positives (NTP or sensitivity): Number of reference transcripts with at least one AllTP.
FSM Redundancy level: Number of FSM divided by the number of FSM reference transcripts
Coverage by long reads: % of the transcript length covered by the supporting reads

ISM transcripts (having a reference transcript with the same junction chain but 3' and/or 5' junctions are missing)
Same metrics as for FMS. Additionally
Longest junction chain (%): Number of junctions in ISM divided by the number of junctions in the matched reference
Intron retention level (IR-ISM): Number of IR transcripts within the ISM category

NIC transcripts (at least one novel junction with known donor and acceptor sites)
Illumina Junction support: % novel junctions with Illumina reads junction support
Illumina NIC support: % NIC transcripts with all novel junctions supported by Illumina reads
% Novel junctions: Distribution of the number of novel NIC junctions per NIC transcript

Supported NIC: % NIC transcripts with reference, CAGE or polyA support, and Illumina novel junction support
Intron retention level (IR-NIC): Number of IR transcripts within the NIC category
Longest junction chain (%): Longest chain of known junctions in the NIC transcript divided by its total number of junctions

NNC transcripts (at least one novel junction with a novel donor or acceptor site)
Same metrics as for NIC. Additionally
Non-canonical splice junctions (ncSJ): % of non-canonical junctions over the total number of *NNC novel junctions*. Canonical junctions are GT/AG, GC/AG, AT/AC.
NNC-non canonical (NIC-nc): % of NNC transcripts with at least one non-canonical junction

For all FSM, ISM, NIC, and NNC, additionally
Count: Number of instances
Distribution of distances to TSS or TTS of reference transcripts.
% of exact matches to reference TTS and TSS (0 nts deviation)
Level of RT-switching incidence.

For all other SQANTI categories
Count: Number of instances
Exon number
Intra-priming evidence.

Evaluation by *bona fide* GENCODE transcripts
Detection: % of *bona fide* GENCODE transcripts detected by at least one FSM, ISM, NIC, NNC match
Sensitivity: Number of FSM with 3' and 5' ends within 50 nts of annotated TSS or TTS divided by the number of *bona fide* GENCODE transcripts
Coverage by long reads: % of the transcript length covered by the supporting reads

Experimental validation
A number of novel transcripts detected by all or most pipelines, as well as pipeline-specific transcripts will be selected for validation by PCR. We will evaluate:

a. Novel junctions
b. Novel combinations of exons

Challenge 2 Evaluation: Transcript isoforms quantification

Evaluation of quantification
Consistency in expression values: We will evaluate the correlation among replicates at different levels of expression.
Accuracy at spike-ins. Estimates of expression levels will be done based on SIRVs and ERCC data
Prediction of fold change. Fold change at the gene and transcript isoform-level between the H1:H1-DE cell line mix versus WTC11. *qPCR* of a selected number of transcript models will be performed.

Challenge 3 Evaluation: *De-novo* transcript isoform detection without a high-quality genome

Evaluation of transcriptome annotation for Manatee models

Transcript models will be described in terms of:

a. number of transcripts per loci/gene
b. length of the transcript models
c. coding potential

A number of loci will be selected for experimental PCR validation as in the mouse/human data.

Ongoing discussions are in place to enable validation by complementary genomics technologies.