DIVERGENE

FAQs

Tips for Improving Libraries

The Pacific Biosciences instruments perform sequencing in real-time using single-molecule primary (not amplified) material. At this time library inputs include DNA, cDNA, or RNA.
Since the instrument uses primary material it is critical that the samples submitted be in very good condition. Here are some suggestions for submitting the best possible sample and thus ensuring the best possible outcome.
As always, please send us your questions, suggestions, and feedback - this is an evolving page. Email us at support@divergene.com.

High Quality DNA (DNA Samples are submitted as GENOMIC DNA, AMPLICON DNA or cDNA.):

- Genomic DNA should be high molecular weight and free of nicks.
- Amplicon DNA should be the result of log phase amplification (don't max out those cycles and reagents!). PCR products should be clean, without non-specific products or multiple bands.
- The cDNA samples should be generated without nicks or small fragments (we recommend the Evrogen Mint-2 Kit or similar style first strand synthesis)
- For AAV material, Divergene recommends the use of the Invitrogen PureLink™ Viral RNA/DNA Mini Kit-12280050 or Takara AAV proPurification Kit (midi)-all serotypes-6675.
- Nicks can be repaired using the PreCR Repair Kit. Please do not use this if you are submitting for base modification detection.
- Samples should be in EB or water.
- Allow DNA to air-dry. Do not heat if drying in a speed-vac.

Divergene offers conversion to cDNA (through the MBCL) if you prefer to submit RNA. If you desire Divergene to perform the DNA/RNA extraction, please contact us to discuss the project ahead of time.

Submit sufficient material for a high-quality build. Current recommendations are:

- AAV PacBio sequencing library requires >500ng AAV DNA per sample in ≤20 ul of EB or water
- Iso-Seq™ (strand-specific RNA-Seq) requires 1-2.5 µg Total RNA
- A cDNA library of 500bp-1.5kb requires 400-800 ng of cDNA
- A long-insert or Hi-Fi genomic library of 1.5-<30kb requires 2-5 µg of genomic DNA
- A long-insert genomic library >=30kb requires 10 µg of genomic DNA
- An amplicon library requires 200-300 ng of amplicon DNA

Request enough SMRTcells for the sequencing depth you need:

- If you are unsure about depth, please discuss your project with Divergene
- Libraries can be barcoded and mixed, then data demultiplexed by Divergene, by two methods: either "internally" using Pacbio's Universal F/R Primers during amplicon prep or with barcoded adapters during library prep. This is recommended for cost-savings if you have multiple libraries that do not require a full SMRTcell of coverage.
- Please note that the predicted number of cells needed is based on the amount of **target** DNA. If the sample is contaminated with DNA from other organisms, more cells will be needed to reach your target depth.

Performance Issues -

Things which can cause libraries to perform badly (or not at all) on a PacBio Run

---

Nicked or Damaged DNA:

- Minimize Free-Thaw Cycles
- Do not expose to pH extremes
- Avoid vigorous vortexing or syringing
- Do not expose to prolonged high temperatures
- Do not expose to intercalating dyes or UV radiation

DNA with quality issues. Try to make sure that your DNA does not:

- Contain insoluble material or contain RNA
- Have an OD260/280 outside the range of 1.8 to 2.0
- Contain EDTA or other chelating agents
- Contain divalent cations like Magnesium or Manganese or Calcium
- Contain denaturants like guanidine or phenol
- Contain detergents like SDS, Triton, Trizol reagent, etc.
- Contain contaminating material from starting tissue like heme from blood

RNA with quality issues. Try to make sure that your RNA does not:

- Contain insoluble material or contain DNA
- Have an OD260/280 outside the range of 2.0 to 2.2
- Have an OD260/230 outside the range of 1.8 to 2.1
- Contain denaturants like guanidine or phenol
- Contain detergents like SDS, Triton, Trizol reagent, etc.
- Contain contaminating material from starting tissue like heme from blood

Sample Submission Process

Please email the Divergene to request the most recent version of our service ticket.

Material should be submitted at the requested concentration listed on the service ticket for your desired library type. Amplicons should be in **15-20 ul of EB or water**, and genomic material is preferred in **15-50 ul of EB or water**. The outcome of libraries made from material submitted below the requested concentration and quality standards may not be guaranteed by the Core.

Please use 1.5mL tubes, 0.5-0.65mL tubes, or PCR strips that are labeled on the top of the tube. **Do not submit single PCR tubes**, and avoid long labeling strips that would prevent the tube(s) from fitting in a centrifuge.

In the data pipeline, characters used for sample names are limited to alphanumerics, "-", or "_".

Please note that a sample submitted for sequencing is entered in the workflow and queue, which engages time and services. If you want the QC analysis only for your own information and are not certain you wish to move forward, please use the Fragment Analyzer or other QC methods (e.g. Topo Cloning, Gel Purification).

Material sent for sequencing should have **one completed ticket PER SAMPLE MIX**, and an Index List form if applicable. Electronic files are always welcome in addition to paper tickets. Typically we suggest that you send indexed libraries separately and let us do the mixing by molarity based on our internal QC. This helps with any potential troubleshooting if some samples in the mix underperform. However, if you prefer to premix your samples or choose a different mixing method (such as by volume), you are welcome to.

> Divergene can demultiplex your library mix after the run if you have used PacBio internal barcode sequences such as the Barcodes Universal F/R Primers or if indexed adapters are used to build the library. If this applies to you, please inform us by filling out the relevant section on the service ticket and sending us a list of sample names vs. barcodes.

Our shipping address for samples:
Drs. Ellen Kittler / Maria L. Zapp
PBCE
222 Maple Avenue
Reed Rose Gordon Building, Room 141
Shrewsbury, MA 01545

Sequencing Process

The Core Lab staff will QC your submitted sample and enter it in the queue. The Core will email you a copy of the QC results upon request. If you then choose not to move forward with sequencing, the Core will bill for the QC and processing to recoup costs. Samples will be processed on a first-come, first-served basis.

Data Analysis

After the data collection is complete, samples enter the bioinformatics pipeline. The time required to complete this varies based on the analyses requested as well as the state of the High Performance Computing Cluster.

If you would like to consult with an applications specialist, please email us to get on the schedule.

IF THERE IS A PROBLEM WITH YOUR RUN OR QUESTIONS ABOUT YOUR DATA, please let us know as soon as possible. The run metrics and machine files cannot be held for long because they are extremely large. The sooner we know there is a problem, the more likely we'll be able to help easily. In order to help sort out a problem we may ask a lot of questions and details about your sample and your analysis methods. This information is required in order for us to engage the tech support systems provided by our vendors.

Please note that if the control library on the run failed in any way, the SMRTcell is rerun. If your data is delivered to you, then the controls all passed spec., which means that the instrument, reagents, and chemistry functioned properly. That leaves us with investigating other issues including sample design, sizing, indexing, degraded DNA, as well as computational and analysis issues with the pipeline, etc. We'll do whatever we can to get things working, but the more information we have, the better.