Sample Preparation
Frequently Asked Questions (FAQ)

Updated August 2013

Complete Genomics data is for Research Use Only and not for use in the treatment or diagnosis of any human subject. Information, descriptions and specifications in this publication are subject to change without notice.

Copyright © 2013 Complete Genomics Incorporated. All rights reserved.

Complete Genomics Sample Requirements

At this time, Complete Genomics accepts human DNA only. DNA may be extracted from blood, frozen tissue, cell lines, or saliva. Complete Genomics cannot accept DNA that has been treated with whole genome amplification (WGA) or that is derived from formalin-fixed, paraffin-embedded (FFPE) samples. The DNA sample requirements are as follows:

Feature	Sample Requirement
DNA amount, as measured by PicoGreen	Strongly recommended: ≥ 10 µg Recommended: ≥ 5 µg
DNA concentration, as measured by PicoGreen	75 to 300 ng/µl
DNA volume	50 to 200 µl
DNA fragment length	High molecular weight double-stranded genomic DNA, majority over 20 kb
DNA buffer	1x TE, pH 8.0 ▪ 10 mM Tris, 1 mM EDTA ▪ H_2O is not recommended.

A260/A280 ratios should be between 1.8 and 2.0. Values outside of this range suggest that there are impurities in the sample that could affect sequencing performance.

Acceptance criteria are based only on measurements performed by Complete Genomics, and are not based on the amounts reported by the customer. Because there is inherent variability in measurement between sites and users, targeting the minimal amount (3.5 µg) will likely result in a significant number of samples failing to meet acceptance criteria and lead to a delay in sample processing. For this reason, Complete Genomics strongly encourages customers to send additional DNA when available. Unused DNA can be returned after sequencing, upon request.

Extracting and Storing DNA

Which kits are recommended for DNA extraction?

DNA samples extracted using the kits listed here have consistently provided high-quality results. In general, commercial kits are recommended because the associated reagents have been subjected to quality control before use and are not likely to introduce problems.

Complete Genomics strongly recommends following manufacturer's guidelines with respect to the amount of cell/tissue extract loaded per column:

© Complete Genomics, Inc.

- Overloading columns decreases DNA yield and increases the likelihood of producing 'dirty' DNA that will not perform well in the sequencing process.
- Reducing the volume of the column washing solution may result in carryover of contamination, which interferes with the DNA fragmentation and library construction process.
- For many kits, treatment with proteinase K and/or RNase is considered optional. Complete Genomics highly recommends using both of these reagents to remove contaminating proteins and RNA. See What if there is RNA in the sample? and What if there is protein in the sample? for more information.

DNA Extraction Kit (Sample Catalog No.)	Blood	Tissue (including tumors)	Cultured Cells	Saliva
Precision System Science Magtration® System 8Lx (E7002)	✓			
QIAGEN® DNeasy® Blood & Tissue Kit (69504)	✓	✓		
QIAGEN® Gentra® Puregene® Cell Kit (158722)			✓	
QIAGEN® Gentra Puregene Blood Kit (158445)	✓			
QIAGEN® QIAamp® DNA Blood Midi Kit (51183)	✓			
QIAGEN® PAXgene® DNA Blood Kit (761133)	✓			
DNA Genotek Oragene OG-500 kit for collection and prepIT•L2P kit for extraction*				✓

* The prepIT•L2P kit (which is based on ethanol precipitation) and the PD-PR-015 whole sample protocol is recommended for extraction of DNA from saliva samples. Extracting the entire Oragene sample allows for maximum DNA recovery and concentration. At this time, the prepIT•C2D kit (which is column-based) is not recommended because it does not allow for the entire sample to be extracted with a single aliquot.

How can I ensure that my saliva-derived DNA is within the correct concentration range?

If using the recommended prepIT•L2P kit for extraction and PD-PR-015 protocol, the DNA can be eluted in 0.2 to 1 mL of TE solution. To ensure that the DNA is in the correct concentration range for Complete Genomics sequencing, but also fully re-hydrated, 500 µl is recommended for the elution volume. Take care not to over-vortex the sample or to heat it for longer than the hour prescribed in the protocol.

How can I reduce the change of ethanol contamination in my DNA extraction using a column-based protocol?

It is important to make sure that the ethanol is completely removed from the sample. If using a column purification method, spinning twice after the ethanol wash step can help completely dry the column and to ensure removal of excess alcohol prior to elution of the DNA.

What type of ethanol should I use for best results?

Wash buffers for DNA extraction should be made using high-quality "200-proof" or "absolute" ethanol. Some vendors sell versions that are "molecular biology grade" or "biotechnology grade."

© Complete Genomics, Inc.

Do not use denatured alcohol or denatured ethanol as they contain isopropanol, methanol, and/or other solvents that can interfere with the production of high-quality DNA.

How would I know if there is ethanol in my DNA?

Ethanol contamination in DNA can cause the following:

- DNA prep smells odd or reminiscent of ethanol.
- DNA "floats" back out of the well when loading a gel, even if loading dye is added to the sample.
- DNA doesn't freeze well at –20°C.

Do you accept DNA extracted with Phenol-Chloroform?

While we have successfully processed samples from customers who have used phenol-chloroform to extract DNA, we have also observed issues. These issues are likely the result of residual amounts of organic solvents that may interfere with the sample handling. If phenol-chloroform has been used to extract DNA, Complete Genomics highly recommends performing a cleanup step—such is as available in the DNeasy Blood and Tissue Kit—after the extraction and before shipping to Complete Genomics.

Do you accept DNA extracted from formalin-fixed, paraffin-embedded (FFPE) samples?

Complete Genomics does not currently accept FFPE samples.

What if there is RNA in the sample?

Contaminating RNA can result in an overestimation of DNA concentration when measured using UV spectrophotometry such as NanoDrop. The use of a quantitation method that specifically measures double-stranded DNA, such as PicoGreen, should help avoid such overestimation. However, large amounts of contaminating RNA can also result in overestimation of DNA concentration due to high occurrence of RNA secondary structures. See Which kits should I use for DNA quantitation? for more information. Contaminating RNA should not affect sequence quality.

What if there is protein in the sample?

Protein-DNA complexes have slower mobility than pure DNA when running through an agarose gel and might be present as a slower migrating DNA band or doublet running above the primary genomic DNA band. Alternately, protein may appear as an intercalator-stained material visible inside the agarose gel wells, since it hasn't entered the gel. Initial results suggest that the presence of such a slow-migrating band may increase the risk of sequencing problems, including an inability to deliver CNV and SV results and an increase in small variant error rates.

To minimize the risk of sequencing problems, Complete Genomics highly recommends that you provide a replacement sample that does not contain slow-migrating DNA. If there is no replacement available, Complete Genomics recommends removing the protein by proteinase K treatment followed by column-based purification. Gel electrophoresis should be repeated to ensure that the DNA-protein complex band has been eliminated.

What is the effect of high temperature during DNA extraction?

Incubation at temperatures above ~45°C during DNA extraction can lead to the generation of single-stranded DNA. This can have the same effect as protein contamination, leading to increased risk of sequencing problems, including an inability to deliver CNV and SV results and higher small variant error rates. Some protocols recommend incubations with Proteinase K at higher temperatures. In

© Complete Genomics, Inc.

this case, it is important to limit the amount of time the DNA sample spends at the higher temperature, ideally to a maximum of a couple of hours.

Which buffer should I use to store my extracted DNA?

To maintain DNA integrity, Complete Genomics recommends that all DNA samples be provided in 1x TE, pH 8.0. Samples should not be supplied in H_2O.

My DNA is eluted in QIAGEN buffer AE. Is this acceptable?

Yes. Though the solution is 1x TE, pH 9, sequencing has been successful with DNA eluted and stored in this buffer.

My DNA is eluted in QIAGEN Gentra Puregene DNA Hydration solution. Is this acceptable?

Yes. The solution is 1xTE pH 7-8.

My DNA was eluted in water instead of 1xTE pH 8. What should I do?

Tris-EDTA (TE) buffer prevents DNA degradation by inactivating nucleases and maintaining a neutral pH. To prevent DNA degradation during shipment, Complete Genomics recommends adding a small amount of 10x TE to bring the final concentration up to 1x. Keep the concentration requirements in mind when doing so to avoid diluting the DNA too much.

How much of the DNA source is required to obtain sufficient DNA for Sequencing?

If you are planning to isolate DNA from blood, refer to the following article from QIAGEN, which investigates the effects of a combination of factors on final DNA yield for one of its recommended kits, the QIAGEN Gentra Puregene Blood Kit:

http://www.qiagen.com/literature/qiagennews/weeklyarticle/08_05/e11/default.aspx

Significant factors include choice of anticoagulant, storage temperature, storage time, and even the method of thawing samples.

Typical yields of some of the recommended kits are described in the following links:

- QIAGEN DNeasy: http://www.qiagen.com/products/genomicdnastabilizationpurification/dneasytissuesystem/dneasybloodtissuekit.aspx#Tabs=t1
- QIAGEN Gentra PureGene Kits:

  http://www.qiagen.com/literature/render.aspx?id=200380

Follow the commercial instructions of the kits exactly and do not overload the columns in an attempt to get additional DNA. For more information, refer to the specifications of the kit used for extraction.

For DNA collection from saliva samples, 2 mL of saliva mixed with 2 mL of Oragene solution, with a total volume of Oragene/saliva solution of about 4 mL, should result in > 20 µg DNA (using the recommended extraction kit and protocol). Saliva sponges used to collect saliva from infants or young children will yield less DNA, as described in this DNA Genotek collection white paper:

http://www.dnagenotek.com/US/pdf/PD-WP-007.pdf

© Complete Genomics, Inc.

Making Sample Measurements

How do I quantitate my DNA?

Complete Genomics performs several quality checks prior to accepting samples for sequencing. These include determining sample quantity using the Quant-iT™ PicoGreen® dsDNA kit from Invitrogen and determining sample integrity by electrophoresis. Both of these approaches should be used for each sample before the sample is submitted to Complete Genomics. Details on the sample requirements can be found in the Sample Submission Instructions, and details on the Sample QC performed after sample receipt can be found in the Sample Quality Control Protocol.

Which kits should I use for DNA quantitation?

Complete Genomics recommends the following fluorometry-based kits that specifically measure the concentration of double-stranded DNA.

Kit Name	Manufacturer	Catalog #
Quant-iT PicoGreen dsDNA Assay Kit	Invitrogen	P7589, P11496
Qubit Quant-iT™ DNA Assay Kit, High Sensitivity *0.2 - 100 ng*	Invitrogen	Q33120
Qubit dsDNA BR Assay Kit	Invitrogen	Q32850, Q32853
Qubit dsDNA HS Assay kit	Invitrogen	Q32851, Q32854

Note: Using the proper standards to create the standard curve is essential to determining the proper DNA quantity. Complete Genomics strongly recommends using the standards provided by the kits.

Note: Using UV absorbance to quantitate DNA may result in a different quantity measured than the assays in this table.

Note: PicoGreen can be highly sensitive to the detergent CTAB, sometimes used for DNA extraction. If you use this approach, make sure to remove any residual CTAB prior to quantitation.

To mitigate the risk of sample failure due to low quantity, Complete Genomics recommends sending as much extra DNA as possible.

What if I don't have access to a recommended DNA quantitation kit?

All samples received by Complete Genomics will be measured using PicoGreen, and the internal measurements will be used to determine which samples can proceed toward sequencing or require replacement or top off.

The benefit of PicoGreen for DNA measurement is that it specifically measures double-stranded DNA, the template used for Complete Genomics whole genome sequencing. When using alternate approaches such as spectrophotometry, contaminating protein, RNA, or single-stranded DNA will contribute to the measured A260 absorbance and thus result in an overestimation of the amount of DNA present in the sample. Other contaminants can also contribute to the A260 absorbance resulting in overestimation of the DNA concentration.

If there is no access to PicoGreen or another approach for DNA measurement using fluorometry, Complete Genomics strongly recommends sending at least 2 times more DNA than estimated by NanoDrop spectrophotometry or an alternate approach. Sending as much DNA as possible is the best way to ensure that there is sufficient DNA for sample acceptance.
An alternate approach using fluorometry for DNA measurement is the Invitrogen Qubit® fluorometer.

© Complete Genomics, Inc.

Why do PicoGreen and NanoDrop Results Differ?

PicoGreen specifically measures double-stranded DNA and therefore provides a more accurate detection of DNA concentration and amount. NanoDrop spectrophotometry indirectly infers DNA concentration by determining the absorbance of light at 260 nm. Unfortunately, while double-stranded DNA has the highest absorbance at this wavelength, other molecules also absorb light; including protein, RNA, single-stranded DNA, free nucleotides, phenol, and other contaminants. Therefore, NanoDrop spectrophotometry measurements are prone to overestimation of DNA amounts due to the detection of contaminants in the sample.

Can I get information on the Sample QC protocol followed by Complete Genomics?

Yes, the detailed protocol is available in the Sample Quality Control Protocol.

What if I'm not sure about my gel results?

Complete Genomics recommends that you provide a gel image to your Complete Genomics Project Manager prior to shipping your samples. If you are not sure about the compatibility of any DNA samples with our sequencing technology, the Project Manager will provide feedback on whether or not the samples are suitable to be shipped and sequenced.

What should I look for in the gel results?

DNA should be intact, with a single band indicating that it is > 20 kb, and minimal degradation. See the Sample Submission Instructions for examples of gel images.

RNA, observed by two bands of ribosomal RNA, can cause a sample to fail Sample QC due to overestimation of concentration, however, it should not interfere with sequencing. See What if there is RNA in the sample? and Which kits should I use for DNA quantitation for more information.

Some agarose gels reveal a smear or band migrating above the main band of DNA, indicating that the sample contains some DNA with abnormally slow mobility (potentially single stranded DNA or the effect of protein contamination). See What if there is protein in the sample? for more information. This slower migrating material may also result from high-voltage or pH gradients (which are more likely when using TAE buffer) during electrophoresis.

Preparing DNA for Shipment

What if the DNA is not in the proper concentration range?

Complete Genomics does not concentrate or dilute samples that are outside of the required concentration range, as described in the Sample Submission Instructions. The liquid handling systems used for sample and library preparation have not been validated for concentrations and volumes outside of the required ranges.

If the DNA is too concentrated, it can be diluted by adding 1x TE, pH 8.0.

If the DNA is too dilute, recommended options to concentrate the sample include the following:

- Amicon® ultra centrifugal filter units from Millipore
- Agencourt Ampure XP beads
- QIAGEN QIAamp DNA Micro Kit (catalog # 56304, Cleanup of Genomic DNA Protocol)
- Ethanol precipitation, using NaAoc and not NaCl

© Complete Genomics, Inc.

Note that there can be loss of DNA during sample concentration, as the yields are not 100%. For example, the estimated yield for genomic DNA using the QIAamp DNA Micro Kit is 60%. Using a speed vacuum for concentrating samples is not recommended due to the risk of cross-contamination and increasing the salt in the sample and should only be used when 100% yield is absolutely necessary. If taking this approach, take extra care to ensure that the sample is not completely dried, which can also result in reduced yield.

What are the most common reasons that samples fail quality control prior to sequencing?

1. Insufficient amounts of DNA or incorrect concentration. To ensure that samples pass QC, we recommend customers measure sample DNA concentration using the PicoGreen assay as described in the Sample Quality Control Protocol, and ship additional DNA to buffer for any difference in the customer measurement from Complete Genomics' measurement.
2. Partial degradation. DNA must be provided in 1x TE, pH 8.0. To minimize the likelihood that samples will fail Sample QC due to partial or extensive degradation, we recommend resolving all samples on a 0.8% agarose gel and sharing the results with your Project Manager prior to shipping samples.
3. Improper DNA storage and shipment. Ensure that sample wells are tightly sealed and that the plate is pre-frozen before placing it in the shipping container. Maintain DNA in proper storage conditions (in a -80˚ C freezer) until ready to ship and then ship securely in sufficient dry ice to ensure that the sample does not thaw in transit.

Filling Out the Sample Manifest

What is a Sample Manifest?

The Sample Manifest is a Microsoft Excel form that provides all of the details of the samples that are being shipped on a plate. Each sample plate is matched with one sample manifest, and the manifest includes all of the information for all of the samples on the plate.

Where do I get a Sample Manifest?

The Complete Genomics Project Manager will provide you with a Sample Manifest. If you have previously shipped samples to Complete Genomics, please check with the Project Manager as there may be an updated version of the Sample Manifest available for your next project.

Where do I get information on how to fill out the Sample Manifest?

Details on how to fill out the Sample Manifest are provided in the Sample Submission Instructions. If you have any questions on how to fill out the Sample Manifest, please contact your Project Manager.

What is the difference between a Top-off Sample and a Replacement Sample?

Top-off samples supplement previously submitted sample that failed sample QC due to inadequate amounts of DNA, or to adjust the concentration of the DNA sample into the acceptable range. Each top-off sample must be identical to the sample that failed Sample QC because they will be mixed together.

Replacement samples are sent to be full replacements of failed samples; they may be the same as the failed sample or they may be completely different samples.

© Complete Genomics, Inc.

Can I ship new samples and replacement samples on the same plate?

All top-off samples must be sent on a separate plate from either new samples or replacement samples. This is because they will be taken through our top-off protocol, which includes confirmation that the two samples to be mixed are identical, before starting our standard Sample QC.

Replacement samples and new samples can be on the same plate. See What is the difference between a Top-off Sample and a Replacement Sample to understand the difference between these sample types.

© Complete Genomics, Inc.