

Introduction This document describes the file `metadata.xml`, which is produced by the primary analysis pipeline. This file includes additional information related to **one** movie acquisition.

The latest version of this documentation is available from the PacBio® Developer's Network, at <http://www.pacbiodevnet.com>.

Metadata The `metadata.xml` file contains instrument information related to the conditions present during a movie acquisition. It also provides key pieces of information for integration with primary and secondary analysis. This file is associated with **one** movie.

Name	Type	Description
InstCtrlVer	String	Instrument control software version.
SigProcVer	String	Signal processing software version.
Run	Run	Container for instrument run-related information.
Movie	Movie	Container for movie-related information.
Sample	Sample	Container for sample-related information.
InstrumentId	String	Unique id assigned by Pacific Biosciences.
InstrumentName	String	Name assigned by the customer.
CollectionProtocol	String	Defines the collection workflow (robotic movement, movie acquisition and so on) for a particular cell.
CollectionNumber	Integer	Collection number for this plate/well. Reagents from one well can be distributed to more than one cell.
CellIndex	Integer	The index of this particular cell within the 8Pac.
SetNumber	Unsigned short	A set defines the group of ZMWs being illuminated. A cell can have up to 2 sets. The default value is 1. The value is 0 if the set is unknown.
EightPac	SupplyKit	8Pac supply information.
TemplatePrep	SupplyKitTemplate	Defines the template (sample) prep kit used for this experiment. Can be used to get back to the primary and adapter used.
BindingKit	SupplyKitBinding	Binding kit supply information.
SequencingKit	SupplyKitSequencing	Sequencing kit supply information.
ReagentTube0	SupplyKitSequencing	Information on reagent tube 0.
ReagentTube1	SupplyKitSequencing	Information on reagent tube 1.
Primary	Primary	Primary analysis-specific information.
Secondary	Secondary	Secondary analysis-specific information.
Custom	Custom	A set of custom key-value pairs, set by a user when creating a run definition.

Name	Type	Description
ExpirationData	ExpirationData	Information on whether consumables were expired. (Optional)

Run This is a container for information related to an instrument run. A **run** can contain multiple cells, wells, and movies.

Name	Type	Description
Name	String	Internal assigned name for a run, which consists of multiple wells. There is no constraint on the uniqueness of this data.
CreatedBy	String	Who created the run.
WhenCreated	DateTime	Date and time when the run was created in the system.
StartedBy	String	Who started the run. This can be different from who created the run.
WhenStarted	DateTime	Date and time when the run was started.

Movie This is a container for the movie-related information. A **movie** corresponds to one acquisition for a cell, one set and one strobe.

Name	Type	Description
WhenStarted	DateTime	Date and time when this movie acquisition started.
DurationInSec	Integer	The length of the movie acquisition, in seconds.
Number	Integer	The number of this movie within the set. This is a unique value when combined with SetNumber .

ExpirationData This is a container for information on whether consumables were expired.

Name	Type	Description
EightPacPastExpiration	Integer	8Pac: The number of days past expiration, if at all.
ReagentKitPastExpiration	Integer	Reagent Kit: The number of days past expiration, if at all.
ReagentTube0PastExpiration	Integer	Reagent Tube 0: The number of days past expiration, if at all.
ReagentTube1PastExpiration	Integer	Reagent Tube 1: The number of days past expiration, if at all.

Sample This is a container for the sample-related data.

Name	Type	Description
Name	String	Name of the sample from this well, set by the user.
PlateId	String	The ID of the sample plate.

Name	Type	Description
WellName	String	Identifies which well on the plate this sample came from.
Comments	String	Optional user-supplied comments about the sample.
Concentration	Double	Sample Input concentration.
SampleReuseEnabled	Boolean	Whether or not Complex Reuse is enabled for this well.
StageHotstartEnabled	Boolean	Whether or not Stage Start is enabled for this well.
UseCount	Integer	Count of usages for this batch of complex.
DNAControlComplex	String	Indicates what type of control was used, if any.

Primary This is a container for the primary analysis-related data.

Name	Type	Description
Protocol	String	Name of the primary analysis protocol.
ConfigFileName	String	Name of the primary analysis configuration file.
ResultsFolder	String	Reserved for use by Pacific Biosciences.
CollectionPathUri	Any URI	User-specified location where the results should be copied after analysis is completed.
CollectionFileCopy	PapOutputFile	Defines the set of files to be copied to the CollectionPathUri. This is a list of available file output types from primary output that can be copied out to the CollectionPathUri. Can be one or more of Trace , Pulse , Base , Fasta , and Fastq .

Secondary This is a container for the secondary analysis-related data.

Name	Type	Description
ProtocolName	String	The secondary analysis protocol name specified in the run definition. Ignored by secondary analysis.
Protocol Parameters	String	The secondary analysis parameters specified in the run definition. Ignored by secondary analysis.
CellCountInJob	Integer	The number of cells in this secondary analysis job, identified by the secondary analysis parameter <code>JobName</code> . Supports automated secondary analysis.

Custom This is a set of key-value pairs specified by the user via the run input mechanism. Uniqueness of keys is **not** enforced, and so this may contain duplicate keys.

Name	Type	Description
KeyValue	KeyValue	Up to six user-defined fields specified in RS Remote.

KeyValue This is one custom, possibly non-unique, key-value pair.

Name	Type	Description
Key	String	Key (attribute) and Value (element content).
Label	String	Optional field labeling the key value pair.

SupplyKit This is generic representation of a supply kit.

Name	Type	Description
Name	String	The nickname of this supply kit, if any.
PartNumber	String	The part number.
LotNumber	String	The lot number
Barcode	String	For use by Pacific Biosciences for tracking purposes.
ExpirationDate	Date	The expiration date for the product.

SupplyKitSequencing This is a more specific template kit representation, which includes SupplyKit fields.

Name	Type	Description
Protocol	String	The reagent-mixing protocol used.

SupplyKitBinding This is a more specific binding kit representation, which includes SupplyKit fields.

Name	Type	Description
Control	String	Deprecated; kept for migration purposes only and will be ignored. Use Sample.DNAControlComplex instead.
IsControlUsed	Boolean	Deprecated; kept for migration purposes only and will be ignored. Use Sample.DNAControlComplex instead.

SupplyKitTemplate This is a more specific template kit representation, which includes SupplyKit fields.

Name	Type	Description
AdapterSequence	String	Adapter DNA sequence. Can have up to 2.
InsertSize	Integer	Approximate size of insert.

For Research Use Only. Not for use in diagnostic procedures. © Copyright 2010 - 2013, Pacific Biosciences of California, Inc. All rights reserved. Information in this document is subject to change without notice. Pacific Biosciences assumes no responsibility for any errors or omissions in this document. Certain notices, terms, conditions and/or use restrictions may pertain to your use of Pacific Biosciences products and/or third party products. Please refer to the applicable Pacific Biosciences Terms and Conditions of Sale and the applicable license terms at <http://www.pacificbiosciences.com/licenses.html>.

Pacific Biosciences, the Pacific Biosciences logo, PacBio, SMRT and SMRTbell are trademarks of Pacific Biosciences in the United States and/or certain other countries. All other trademarks are the sole property of their respective owners.

P/N 001-141-374-02