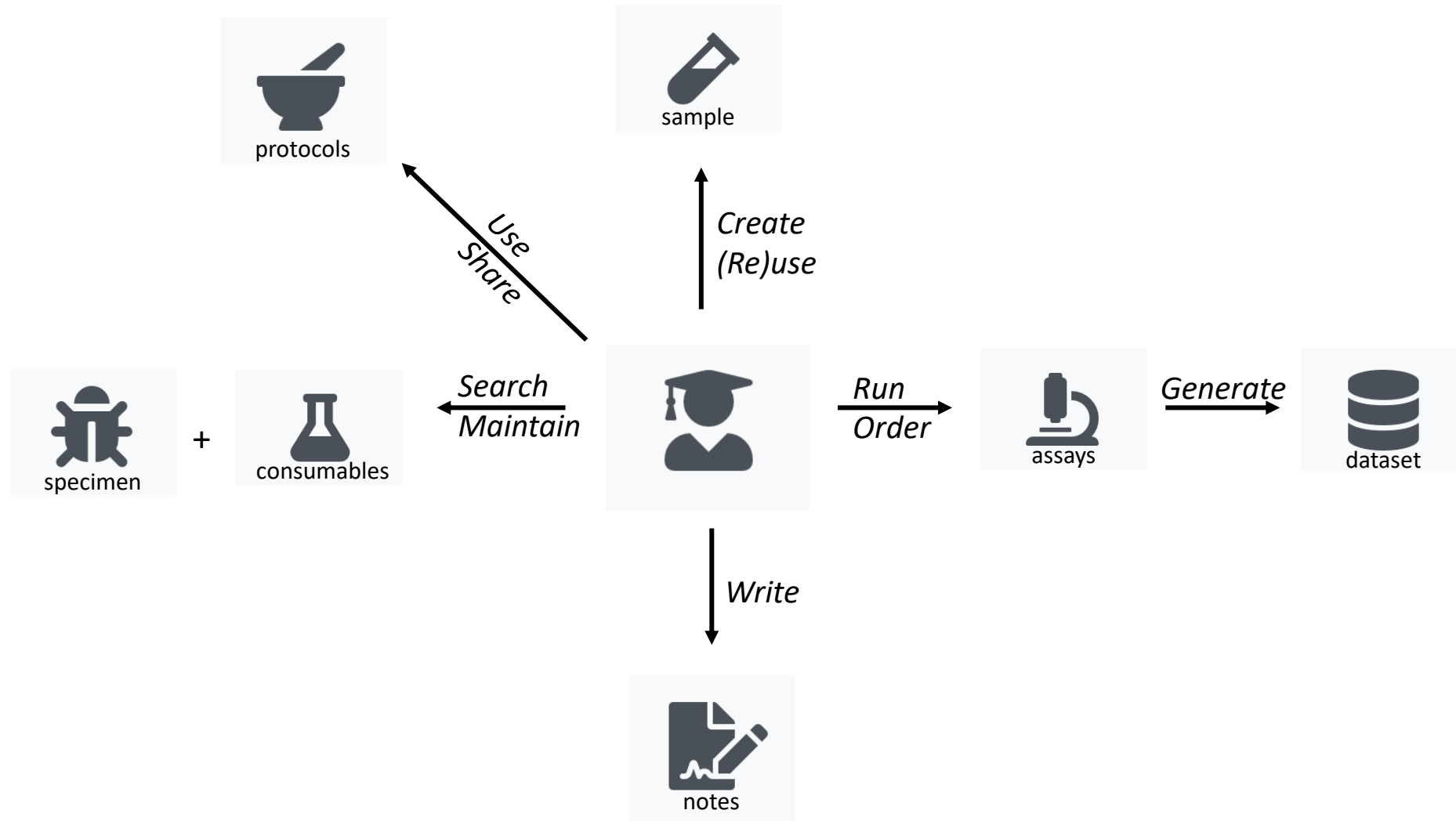


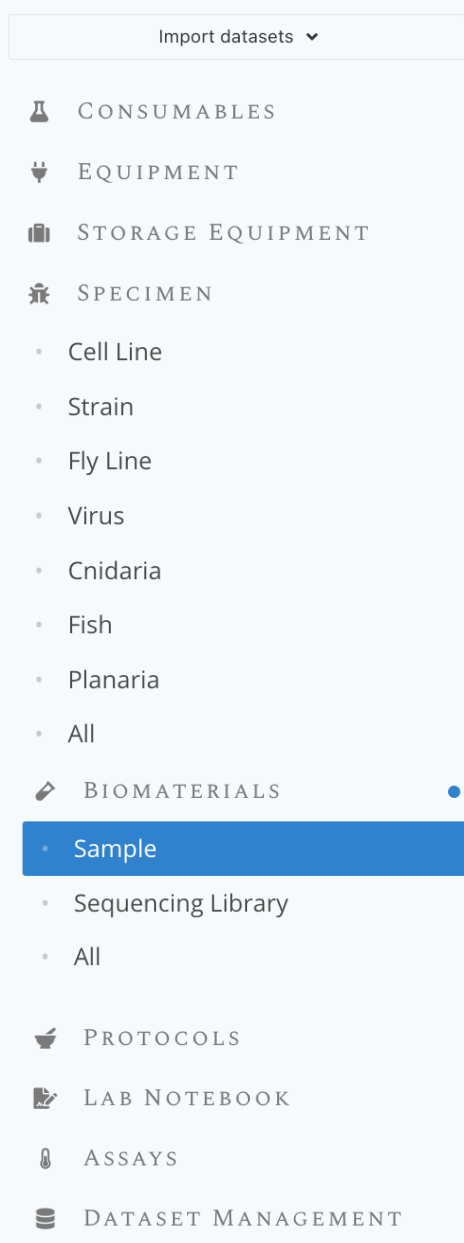
A photograph of a yellow concrete ledge with two metal handrails, set against a bright blue background. The handrails are made of a dark metal and are mounted on the ledge. The concrete has some visible cracks and texture. The blue background is a solid, vibrant color.

STOCKS Training: Sample Management
Charles Girardot
Genome Biology Computational Support (GBCS)

Samples are consumed and created everyday to generate data



The sample management module



- **Specimen:** cell line, fly lines, strains...
 - The living biomaterial from which your samples are coming from
 - **Parent-child relationship**
 - More **specimen types** can be easily **added**
 - **Additional properties** can easily be **added**
- **Biomaterial (aka sample):**
 - **Consumed** and **produced** during your **experiments**
 - The **Sample model** is the **base** model fitted for most needs
 - **Sequencing Library** extends the base sample and add options specific to **sequencing assays**
 - New **sample sub-types** require **addition of code**
 - Missing properties can be added to existing types

Specimen are very similar to the lab inventory collections

petit fish	
<a>Edit <a>Clone <a>Copy URL <a>Add to order <a>Delete	
Item details	
ID	8017e91f-5aee-4b6e-bd07-356fb4d4dbee
Organism	Medaka (<i>Oryzias latipes</i>)
Risk Group	RG1 (Biological agent that are not associated with disease in healthy adult humans or animals)
Code	FM0045
Tank Number	T0045
Father	<a>papa fish
Mother	<a>maman fish
Genotype	fl/+;cre/-
Genetic Background	balb/c
Genetic Modification Technique	CRISPR
Transgenic Insert Plasmid	<a>cre KO plasmid
Transgenic Insert Description	see plasmid
Transgenic Material Donor	Bacillus subtilis
Transgenic Donor Risk Group	RG1 (Biological agent that are not associated with disease in healthy adult humans or animals)
Description	<div style="border: 1px solid #ccc; padding: 5px;">this is how I did it and other cool information</div>
Stock locations	None.
Status	AVAILABLE
Maintained	<input checked="" type="checkbox"/> yes
Responsible Person	<a>girardot
Produced	
Owner	<a>girardot
Owned by Group	Computational Support Genome Biology
Modified	2020-12-17 13:13:47
Created	2020-12-17 11:04:40
Permissions	

➤ Similar features as for the consumables / equipment collections

Properties to parents can be defined when relevant

GMO-related properties to support (German) regulation needs

The Sample biomaterial has the usual properties

Primary project & organism are mandatory

Storage Location(s) available

Sharing & ownership

BIOMATERIAL - GENERICSAMPLE

Organoid S03_CHIR_5microM

ID: dd3a2a98-a0eb-4f09-b557-a9de6aee6d69

Name: Organoid S03_CHIR_5microM

Project: Topology in early Dm embryo (Gaby Project, Furlong Lab)

Status: CREATED | Sample QC: PASSED

Organism: Zebrafish (Danio rerio) | Material Type: organoid

Description: —

Storage Location: 1. Plate 34:A12, Furlong Freezer 1 (-20)

Workflow: —

Sample Lineage

Derives from: Specimen

Specimen: MZ Fz7a/b

Children: —

Is Control Sample? (✗ no)

Is Single Cell Sample? (✓ yes)

Single Cell Type: normal cell | Cell Number: 1200

Screen Plate Position (✗ no)

Link to ELN

Used in Experiments: —

Produced by Experiment: —

Sharing

Ownership and Lifecycle

Annotations (5)

Organism: Zebrafish (Danio rerio)

Growth Media: Avenger++

Sex: female

OrganismPart: Heart

Age: 3 days

Attachments (0)

Drag & Drop your files or Browse

Notes (0)

Write a note | Add

None.

Lineage

Samples (1) | Specimen (1)

MZ Fz7a/b → Organoid S03_CHIR_5microM

Annotations

Attachments

Notes

The Sample biomaterial should fit most use-cases

BIOMATERIAL - GENERICSAMPLE

Organoid S03_CHIR_5microM

ID	dd3a2a98-a0eb-4f09-b557-a9de6aee6d69		
Name	Organoid S03_CHIR_5microM		
Project	Topology in early Dm embryo (Gaby Project, Furlong Lab)		
Status	CREATED	Sample QC	PASSED
Organism	Zebrafish (Danio rerio)	Material Type	organoid
Description	—		
Storage Location	1. Plate 34:A12, Furlong Freezer 1 (-20)		
Workflow	—		
Sample Lineage			
Derives from	Specimen		
Specimen	MZ Fz7a/b		
Children	—		
Is Control Sample?	(X no)		🔕
Is Single Cell Sample?	(✓ yes)		🔕
Single Cell Type	normal cell	Cell Number	1200
Screen Plate Position	(X no)		🔕
Link to ELN			
Used in Experiments	—		
Produced by Experiment	—		
Sharing			
Ownership and Lifecycle			

Activable sections

➤ Designed for supporting array-based screens

Support for control samples

Is Control Sample? (✓ yes) 🔕

Control Name Control Type

Support for plated samples

Screen Plate Position (✓ yes) 🔕

Plate Name	Plate 23	Plate Code	2368491
Plate Type	96well	Plate Position	B12

Samples can have parent sample(s) or specimen

ID [▲] 7f027e5e-9cdb-41d0-bae3-15f68f76f29c

Name * in_1

Project *
 Test Jelle

Status * CREATED Sample QC

Organism * Zebrafish (Danio rerio) Material Type

Description

Storage Location [+ Add a stock location](#)

Workflow
 ...or add an individual protocol

Sample Lineage

Derives from Specimen Sample None

Specimen
 tdgf1 tz257 ; Tg (gsc:GFP-CAAX)

Children [▲] salib2

Is Control Sample? (X no)

Is Single Cell Sample? (X no)

Screen Plate Position (X no)

Link to ELN

Used in Experiments [▲] [Analysis of ATACi samples up to Jan 2020 Batch](#)
 [Nuclei-extraction-and-counting—2019-09-13-13:27:39\(deleted\)](#)

Produced by Experiment [▲] —

Sharing

Annotations 0

—

Attachments 0

Powered by PQINA

Notes 0

—

Lineage Graph

● Samples (2) ■ Specimen (1)

Everything

tdgf1 tz257 ; Tg (gsc:GFP-CAAX) → in_1 → salib2

Either:

- One Specimen; or
- One or more Samples

The lineage reflects the complete experimental procedure

ID: 4acba913-8e17-4d75-934c-231d8b0e9819

Name: Embryos 0-1h yw 5 x 100eeq

Project: something-seq development

Status: CREATED | Sample QC: PASSED

Organism: Fly (Drosophila) | Material Type: whole organism (Whole Organism)

Description: —

Storage Location: —

Workflow: —

Sample Lineage

Derives from: None

Children: RNA 0-1h yw 8 X aliquots 1ug/10ul

Is Control Sample? (X no)

Is Single Cell Sample? (X no)

Screen Plate Position (X no)

Link to ELN

Used in Experiments: yw RNA extraction for smth seq

Produced by Experiment: yw embryo collection for RNA extraction

Sharing

Ownership and Lifecycle

Annotations 0

Attachments 0

Notes 0

Write a note [Add]

Lineage Graph

Assays (1) | Datasets (4) | Samples (9)

Everything | Find Element

```
graph LR; A[Embryos 0-1h yw 5 x 100eeq] --> B[RNA 0-1h yw 8 X aliquots 1ug/10ul]; B --> C[RNA 0-1h yw 8 X aliquots 1ug/10ul]; B --> D[RNA 0-1h yw 8 X aliquots 1ug/10ul]; B --> E[RNA 0-1h yw 8 X aliquots 1ug/10ul]; B --> F[RNA 0-1h yw 8 X aliquots 1ug/10ul]; B --> G[RNA 0-1h yw 8 X aliquots 1ug/10ul]; B --> H[RNA 0-1h yw 8 X aliquots 1ug/10ul]; B --> I[RNA 0-1h yw 8 X aliquots 1ug/10ul]; C --> J[RNA 0-1h yw 8 X aliquots 1ug/10ul]; D --> J; E --> J; F --> J; G --> J; H --> J; I --> J; J --> K[RNA 0-1h yw 8 X aliquots 1ug/10ul];
```

Embedded viewer with many display options and name search

- ✓ As many steps as you need
- ✓ Sample Merge allowed

Samples must indicate how they were generated

- Workflow = [Protocol1, Protocol2, ... , ProtocolN]
- Reflects how the sample was generated from the immediate parent

DBal_Hi-C_N1pat_4-8h_Rep1 Edit Clone Create Child Copy URL

ID	59b3dc0c-f6bf-41be-bcb5-31953ffd76ac		
Name	DBal_Hi-C_N1pat_4-8h_Rep1		
Project	emBASE Migration - Furlong Group		
Status	CREATED	Sample QC	—
Organism	Fruit Fly (Drosophila)	Material Type	synthetic DNA (Synthetic DNA)
Description	—		
Storage Location	—		
Workflow	Double Balancer Hi-C/Capture-C Extraction Protocol ↓ Double Balancer Hi-C Library Preparation Protocol		

Annotations 2

emBASE ID: [embase:labeledextract:18261](#)

emBASE URL: http://embase.embl.de/ngs_sequencinglib_edit.phtml?i_ltext=18261

Attachments 0

Drag & Drop your files or [Browse](#)

Powered by PQINA

Notes 0

Write a note Add

➤ Needed at publication time

Samples are consumed and generated by experiments

➤ Samples are connected to your ELN's *experiments*

ID	006cf1d3-48d6-4436-906f-5893e435dfe7		
Name	RNA 0-1h yw 8 X aliquots 1ug/10ul		
Project	Test Jelle		
Status	CREATED	Sample QC	PASSED
Organism	Fly (Drosophila)	Material Type	nuclear RNA (Nuclear RNA)
Description	—		
Storage Location	—		
Workflow	—		
Sample Lineage	▼		
Is Control Sample?	(X no)		
Is Single Cell Sample?	(X no)		
Screen Plate Position	(X no)		
Link to ELN	▶		
Used in Experiments	Xrn1 + something seq # 13 Generation of RNA-seq lib 19/5/2021		
Produced by Experiment	yw RNA extraction for smth seq		
Sharing	▼		
Ownership and Lifecycle	▼		

- Samples may be consumed or “used” in multiple experiments
- Samples are produced by a unique experiment

➤ Connecting samples to experiments happens on the Experiment page

Link up Samples & Experiments with the Sample Editor

EXPERIMENTS - ENTRY

Delete item Copy item URL Turn off edit mode Save Edits Validation errors

Name* ChIP-WB (SICAP) - 2019-07-17 14:57:37 Project* Start typing Status* IN PROGRESS Created 2019-07-17 16:57:37

This field is mandatory.

Edit Insert Format Table

Paragraph B I U List A D Link Image Video

Sonication:

1. Add 300 ul RIPA+PI buffer to 30-60 min nuclei, pipet up and down 5 times (the number of nuclei per SICAP experiment is based on the starting protein)
2. Let nuclei stand on ice for 10 minutes
3. Transfer the lysate to Bioruptor tube. Sonicate 12 cycles 30sec on/30 off
4. If sonicated several batches of nuclei, pool the sonicates in one tube.
5. Centrifuge @12000g 5 min @4C
6. Transfer the supernatant to a new tube(s).

IP:

7. For each sample split 900ul sonicated chromatin (180*10⁶ nuclei) into 4 aliquots (~220 ul each)
8. Adjust the volume in each sample to 900 ul with RIPA+PI buffer.
 - Perform the preclearing step if used stained nuclei:
 - For each precipitation (also remember the mock reactions!) wash 25

ChIP-WB (SICAP) - 2019-07-17 14:57:37

Deleted * false

Summary

Start date

Completed

Estimated completion date 2019-07-24

Frozen * false

Freeze date

Owner * girardot

Permissions

Public item View Edit Delete

Groups

Add group

Computational Support Genome Biology

Users

Add user

girardot

Samples

Edit

Edit samples

Search for samples Create new input samples Create new output samples Save

Edit area

Edit samples in this section and stage them before saving.

Staging area

Samples have to be staged before saving. For editing, unstage them.

Unstage all

Alerts & Actions

Input samples

TPP #8 Rep 2 (Sample) # parent samples: 1

TPP #8 Rep 3 (Sample) # parent samples: 1

TPP #8 Rep 4 (Sample) # parent samples: 1

TPP #8 Rep 4 (Sample) # parent samples: 1

Output samples

Child1 new (Sample) Add parent specimen

Child1 new (Sample) Add parent specimen

Child3 new (Sample) Add parent specimen

Child4 new (Sample) Add parent specimen

Diagram

TPP #8 Rep 2

Child1

TPP #8 Rep 3

TPP #8 Rep 4

Child3

Child4

- Describe the **input** and **output** samples
- **Connect** samples (parent-child)
- **Existing** samples can be used
- **New** samples can be batch created

The Sequencing Library fits the Sequencing Assay needs

➤ The Sequencing Library = the base Sample + sequencing library-specific attributes

ID	e83e08e0-bd8e-4ce3-b4c3-e92e1f137fcd		
Name	Pool17nonfragsample14		
Project	Phenotype-based single-cell transcriptomics reveal new pathways involved in Golgi organization.		
Status	CREATED	Sample QC	PASSED
Organism	Human (Homo sapiens)	Material Type	polyA RNA (PolyA RNA)
Description	—		
Storage Location	—		
Workflow	—		
Sequence Library Preparation			
Library Kit Type	—	Orientation	—
Barcode	CTCTCTAC		
Source	TRANSCRIPTOMIC SINGLE CELL (RNA-seq of coding RNA from single cells or RNA-seq of non coding RNA from single cells)	Strategy	RNA-Seq (RNA-Seq)
Selection	Oligo-dT (enrichment of messenger RNA (mRNA) by hybridization to Oligo-dT)	RT Primer Type	oligo-dT
Library Strand	NA	End Bias	none
Sample Lineage			
Derives from	Sample		
Sample	HeLa GalNac siUSO1 Non-fragmented Plate3		
Children	—		
Is Control Sample? (X no)			
Is Single Cell Sample? (✓ yes)			
Single Cell Type	normal cell	Cell Number	1
Single Cell Library Construction	Smart-seq2	Single Cell Isolation	Other
Single Cell Well Quality	—		
Spike-ins			
Screen Plate Position (X no)			

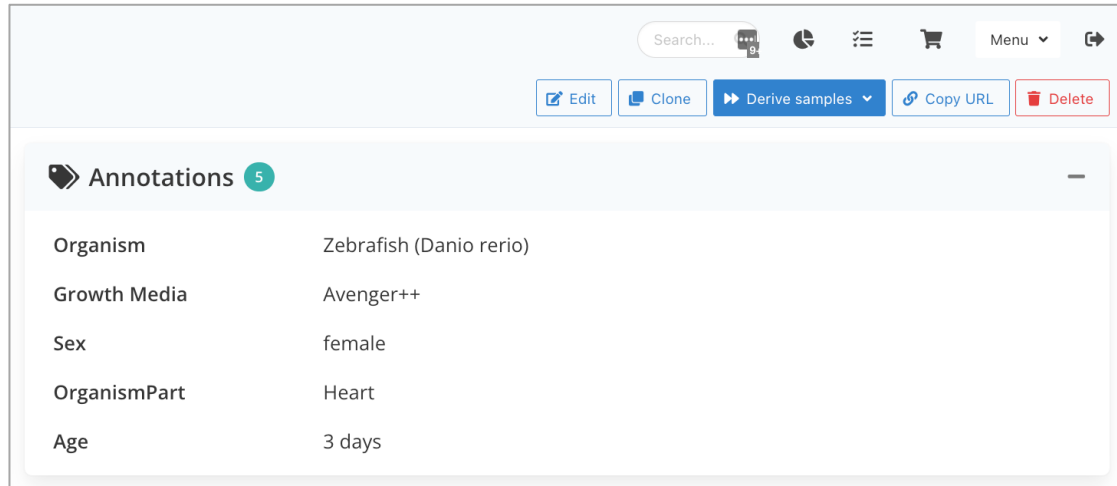
Sequencing library preparation section

- Key aspects of the library preparation procedure
- **Required at submission time**
- Values match those required by public repositories

Single-cell sample section augmented with properties required by public repositories

Additional spike-ins section to reflect common NGS design

Samples must be annotated to be FAIR



The screenshot shows a web interface for sample management. At the top, there is a search bar and several utility icons. Below that, a row of action buttons includes 'Edit', 'Clone', 'Derive samples', 'Copy URL', and 'Delete'. The main content area is titled 'Annotations 5' and contains a table with the following data:

Property	Value
Organism	Zebrafish (Danio rerio)
Growth Media	Avenger++
Sex	female
OrganismPart	Heart
Age	3 days

- Annotations reflect both the project and the sample specificity
- Reflect general sample characteristics e.g. cell line, cell type, organism, sex, genotype, ...
- Define important experimental factors i.e. reflecting the study design
 - How do samples group together e.g. treatment, phenotype, targeted epitope (ChIP) ?
- Biological annotations are mandatory to deposit data in public repositories

Linking samples to experiments and protocols is not redundant

DBal_Hi-C_N1pat_4-8h_Rep1

ID	59b3dc0c-f6bf-41be-bcb5-31953ffd76ac		
Name	DBal_Hi-C_N1pat_4-8h_Rep1		
Project	emBASE Migration - Furlong Group		
Status	CREATED	Sample QC	—
Organism	Fruit Fly (Drosophila)	Material Type	synthetic DNA (Synthetic DNA)
Description	—		
Storage Location	—		
Workflow	Double Balancer Hi-C/Capture-C Extraction Protocol ↓ Double Balancer Hi-C Library Preparation Protocol		
Sample Lineage	▼		
Is Control Sample? (X no)	<input type="radio"/>		
Is Single Cell Sample? (X no)	<input type="radio"/>		
Screen Plate Position (X no)	<input type="radio"/>		
Link to ELN	▶		
Used in Experiments	Xrn1 + something seq # 13 Generation of RNA-seq lib 19/5/2021		
Produced by Experiment	yw RNA extraction for smth seq		
Sharing	▼		
Ownership and Lifecycle	▼		

- Samples can be linked to both protocols and ELN's experiment
 - *Not everyone uses ELN*
- Samples **must** link to protocols
 - At data deposition time, we submit protocols not your ELN's notes
 - In fact we even need a protocol summary (no structured text)
 - Protocols have a *summary* property for this

Workflow templates can be created for common procedures

Single-Strand polyA RNA Illumina Sequencing from ES cells

ID *

Name*

Description

Is Template * yes

Protocols*

- 1. ES_cells_growth_protocol
- 2. Poly(A) mRNA isolation
- 3. cDNA from RNA - SuperScript IV
- 4. mRNA-seq

Modified

➤ Create a workflow template reflecting a chain of protocols often used (e.g. core facilities)

Description

Storage Location [+ Add a stock location](#)

Workflow



Description

Storage Location [+ Add a stock location](#)

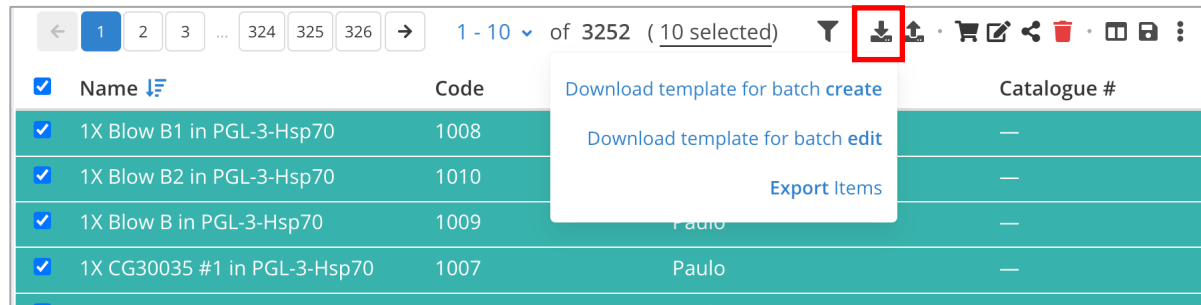
Workflow

Sample Page offers 2 options under workflow

- We used the second one before

Sample Page

Samples can also be batch-created and edited



The screenshot shows a web interface with a table of samples. The table has columns for Name, Code, and Catalogue #. The first four rows are selected. A dropdown menu is open over the 'Download template for batch create' option.

Name	Code	Catalogue #
1X Blow B1 in PGL-3-Hsp70	1008	—
1X Blow B2 in PGL-3-Hsp70	1010	—
1X Blow B in PGL-3-Hsp70	1009	—
1X CG30035 #1 in PGL-3-Hsp70	1007	—



- Works as for any lab inventory collection
- Few additional tricks to know
 - The **ID** of the **Parent(s)** and **Protocol(s)** must be provided, not their name
 - When providing multiple protocols, their order is preserved in the workflow

Sample can be registered in different ways ...and duplicated

1. Manually from the list page
 2. With the Experiment's sample editor
 3. In batch using excel-templates
 4. *At Dataset Registration*
 - similar to the experiment's sample editor
- Samples get duplicated for many reasons
- You create a new one instead of using the existing one
 - Your colleague already created the same sample
 - The sample was existing but not shared with you
 - ...

Getting rid of redundant samples



BIOMATERIAL - GENERICSAMPLE

1 - 10 of 33 (9 selected)

Merge samples

Created	Name	Organism	Workflow	Parents	Children	Description
2022-08-10 15:52:28	sample C	Rabbit				
2022-08-10 15:52:28	sample C	Rabbit				
2022-08-10 15:52:28	sample C	Rabbit				



Merge Samples

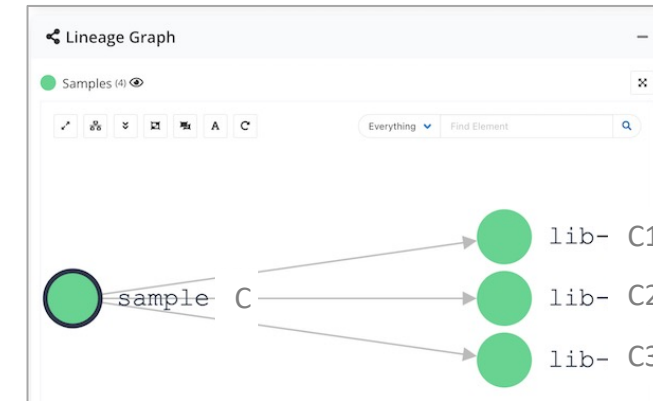
Double check your settings. If mistakes are made, there is no possible rollback to a previous state.

Samples have been grouped by name. Pick one sample to keep for each group before validating.

Merge Group 1

Name	Created	Organism	Workflow	Parents	Children	Description	Material Type	Stock locations	Sample Status	Sample QC	Is control	Keep
sample C	2022-08-10 15:52:28	Rabbit			lib-C1				CREATED		X no	<input checked="" type="checkbox"/>
sample C	2022-08-10 15:52:28	Rabbit			lib-C2				CREATED		X no	<input type="checkbox"/>
sample C	2022-08-10 15:52:28	Rabbit			lib-C3				CREATED		X no	<input type="checkbox"/>

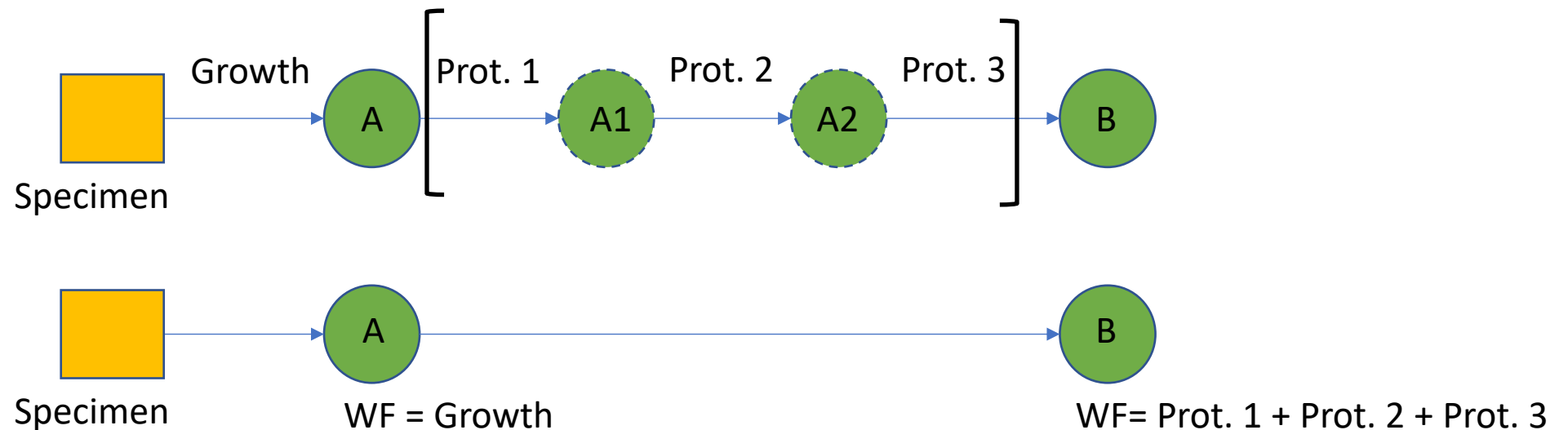
Merging 9 Samples into sample C. The 2 other Samples will be deleted in the process.



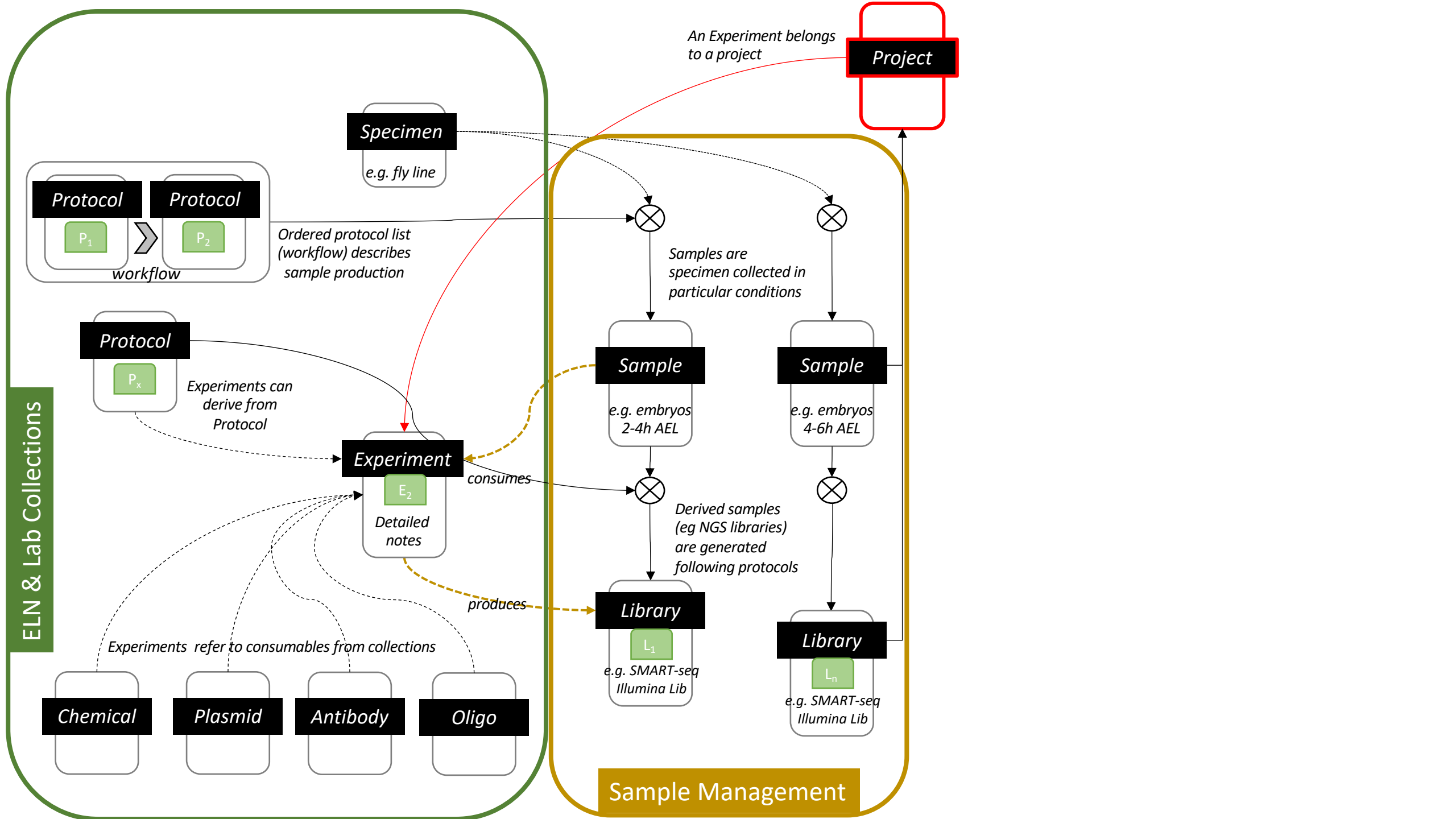
- One sample is kept, the others deleted
- Parent-child relationships are migrated to the kept sample

When should samples be created?

- register your samples as they are generated
 - when taking your notes
- usually not needed to reflect transient samples
 - create samples when found in storage and can be re-used



- To reflect experimental design aspects (parallel processing using split-merge...)



Thank you for your attention



Jelle Scholtalbers



Matthias Monfort



Nayeem Reza

Questions ?