modENCODE
Data Coordination Center

www.modencode.org

modENCODE Scope

- Transcriptome
- Regulatory elements
- Chromosome and chromatin associated elements
- Small and microRNAs
- Origins of replication

modENCODE is a Resource

- DCC: Delivers data to you, the community
- We ask you to tell us what you want
- DCC is listening

~100 Year History of Community Cooperation

- Resource/ Stock Centres (BDGP, CGC, DGRC…)
- Shared reagent resources
- Genome projects (single, multiple)
- Annotation jamborees
- Online resources

- *Drosophila*/ *C. elegans* coordination (modENCODE)

modENCODE in Relation to ENCODE

- Both working on functional elements of DNA
- Exchange of technologies with ENCODE (UCSC)
- modENCODE advantages:
  - can do experimental validation
  - identify functional elements in repetitive sequences
  - Comparisons between worm and fly (and human…)
- This community must set an example for ENCODE

Who are the DCC?

- Principal Investigators
  - Lincoln Stein (PI)
  - Suzi Lewis (coPI)
  - Gos Micklem (coPI)
  - Jim Kent (coPI)
- Wiki and Web Site
  - Sergio Contrino
  - François Guillier
- ENCODE Contacts
  - Kate Rosenbloom
  - Galt Barber
- Data Managers
  - E.O. Stinson (fly)
  - Nicole Washington (fly)
  - Sheldon McKay (worm)
  - Zheng Zha (worm)
- Infrastructure
  - Richard Smith
  - Kim Rutherford
  - Chris Mungall

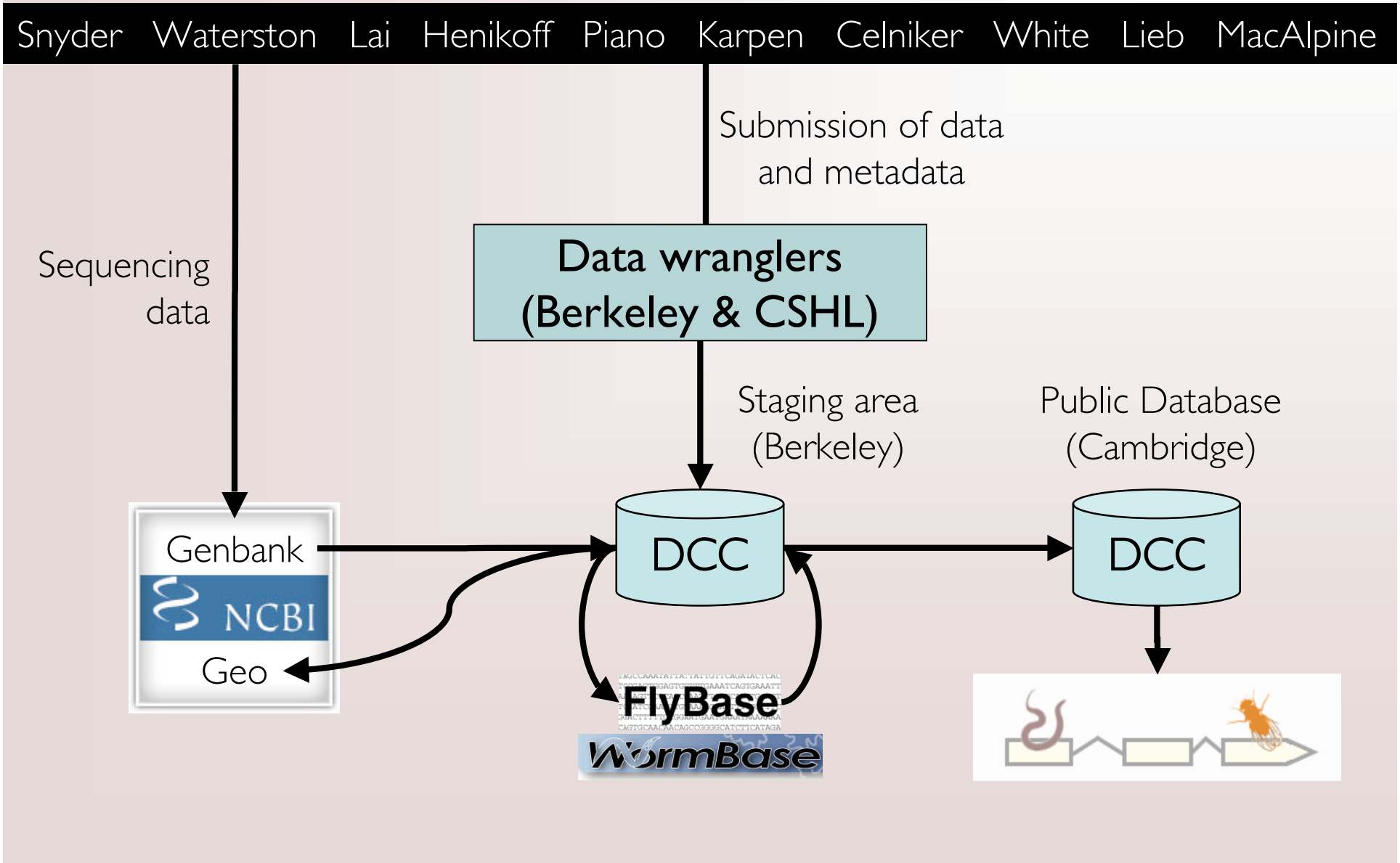
The DCC's Charge

- Provide support for data submission and tracking
- Ensure data integrity and consistency
- Integrate with other information sources
- Make available to the research community

Types of Queries

- Retrieve data by:
  - Protocol type
  - Submitting lab
  - Platform (e.g. Affy Chip set 5.x)
  - Experimental attribute (e.g. Show me the expression array results that use a sliding window > of 30bp)
- Ask useful questions:
  - What antibodies were produced in rabbit that work for ChIP-chip?
  - What is the EST evidence for a particular gene model?
  - What gene models lack cDNA evidence for their TSS?
  - What genomic regions show transfrags that don't have RNA Pol II binding sites within 500bp?

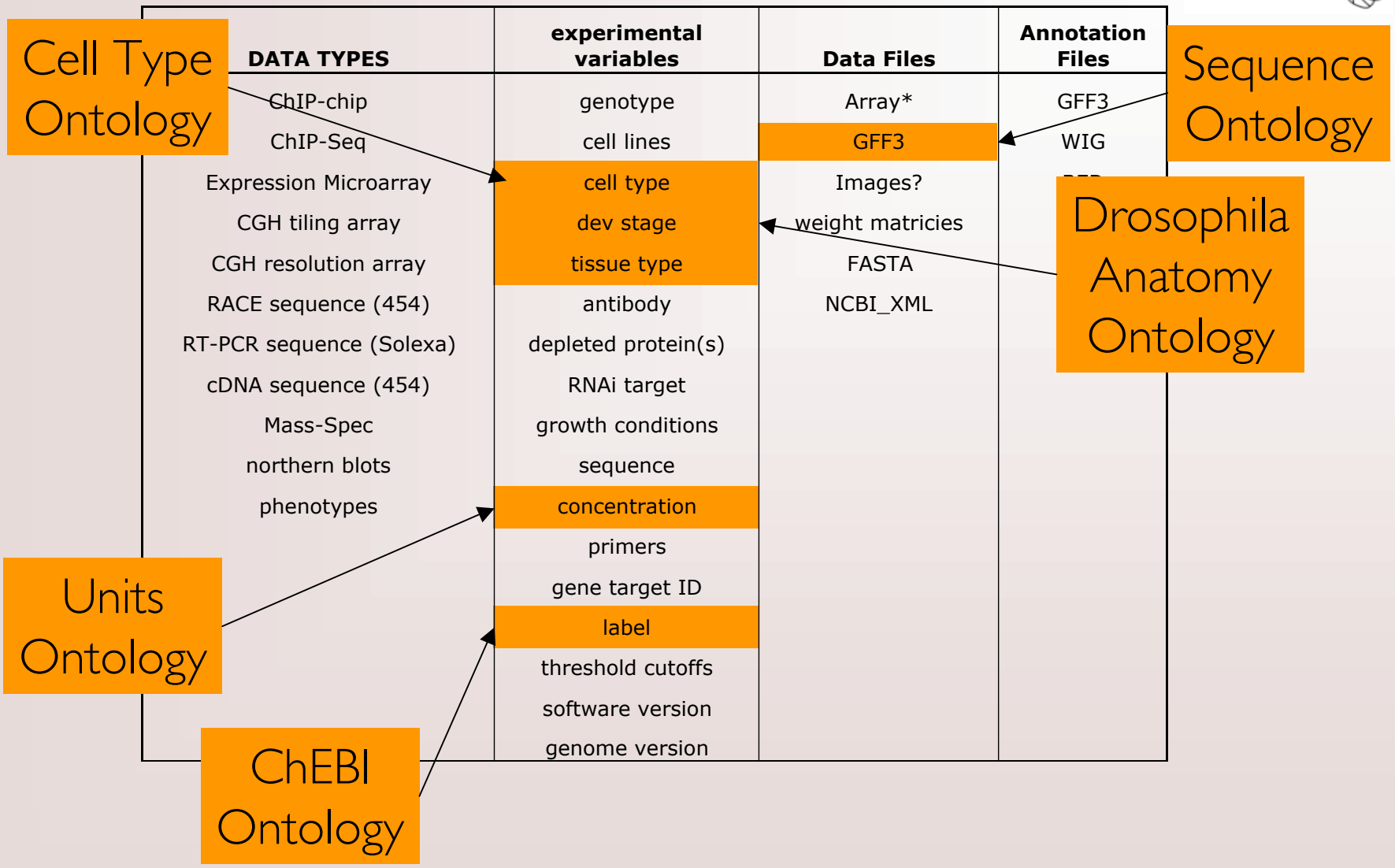
modENCODE & the DCC



Data type summary

DATA TYPES	experimental variables	Data Files	Annotation Files
ChIP-chip	genotype	Array*	GFF3
ChIP-Seq	cell lines	GFF3	WIG
Expression Microarray	cell type	Images?	BED
CGH tiling array	dev stage	weight matricies	
CGH resolution array	tissue type	FASTA	
RACE sequence (454)	antibody	NCBI_XML	
RT-PCR sequence (Solexa)	depleted protein(s)		
cDNA sequence (454)	RNAi target		
Mass-Spec	growth conditions		
northern blots	sequence		
phenotypes	concentration		
	primers		
	gene target ID		
	label		
	threshold cutoffs		
	software version		
	genome version		

Data type summary

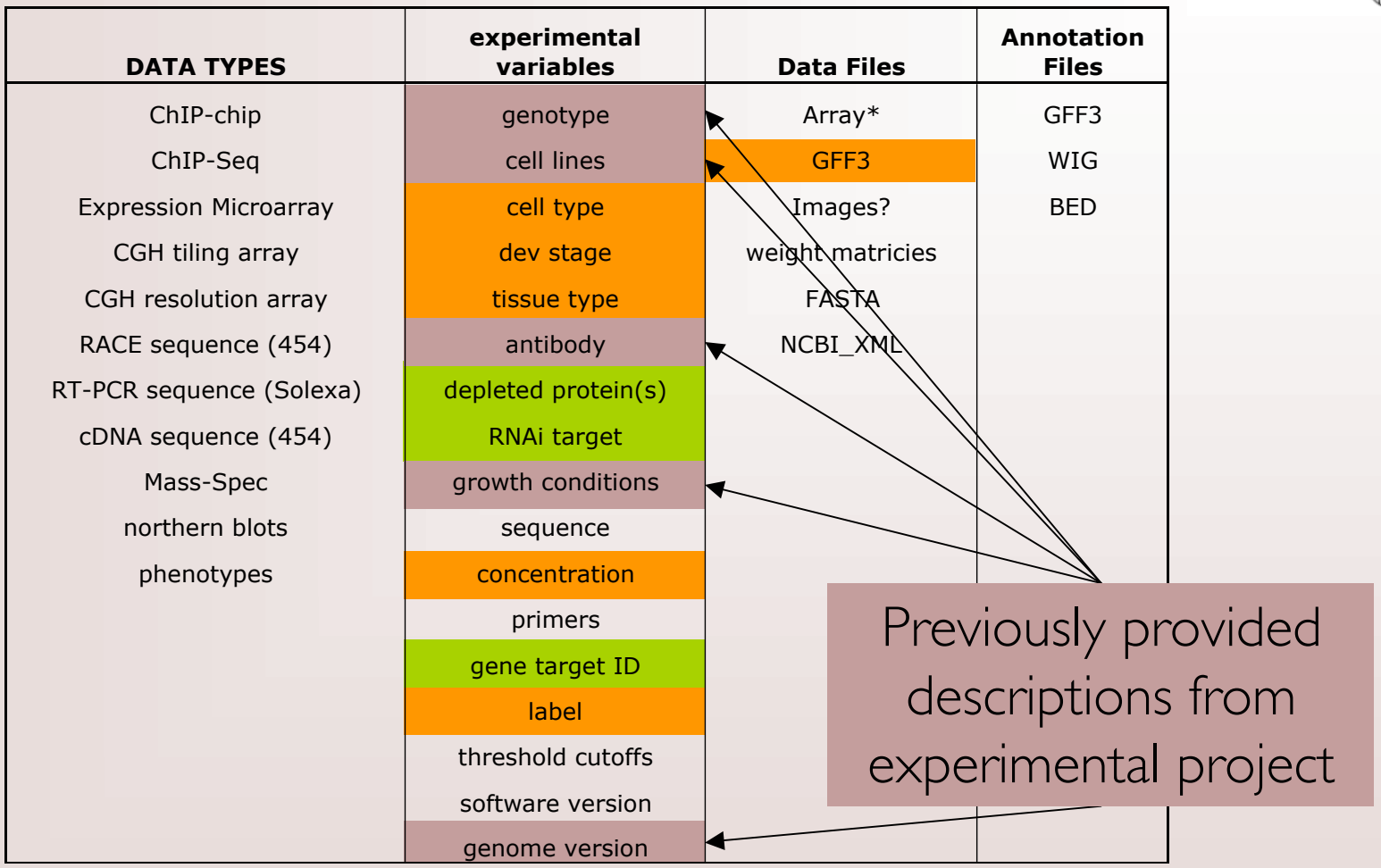


Data type summary

DATA TYPES	experimental variables	Data Files	Annotation Files
ChIP-chip	genotype	Array*	GFF3
ChIP-Seq	cell lines	GFF3	WIG
Expression Microarray	cell type	Images?	BED
CGH tiling array	dev stage	weight matricies	
CGH resolution array	tissue type	FASTA	
RACE sequence (454)	antibody	NCBI_XML	
sequence (Solexa)	depleted protein(s)		
sequence (454)	RNAi target		
Mass-Spec	growth conditions		
northern blots	sequence		
phenotypes	concentration		
	primers		
	gene target ID		
	label		
	threshold cutoffs		
	software version		
	genome version		

Permanent public ID

Data type summary



Biological Investigation Report Tabular Format: BIR-TAB

- A superset of microarray submission formats (MAGE-TAB)
- Extensive support for controlled vocabulary
- Supports any protocol that can be abstracted as an operation with inputs and outputs

Steps in Validation

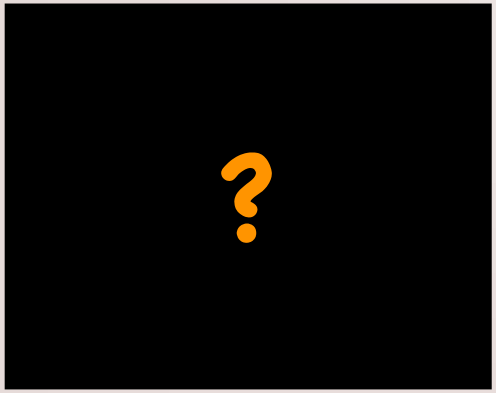
- Syntax checking
- Only terms from specified controlled vocabularies
- External identifiers exist
- Agrees with previously supplied protocol and reagents descriptions
- Check format of data files (currently WIG, GFF3, and BED supported)
- Load into database
- Generate downstream output files

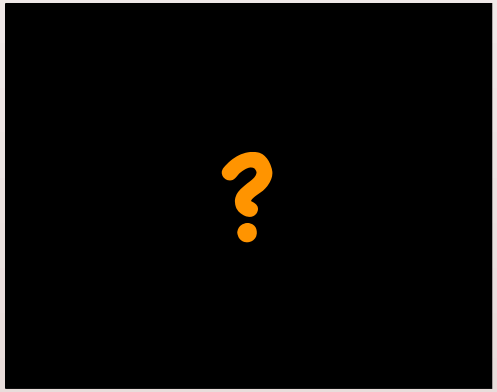
Why Consistency

- Comparison, CompaRison, Comparison
- Without consistency there can be no comparison

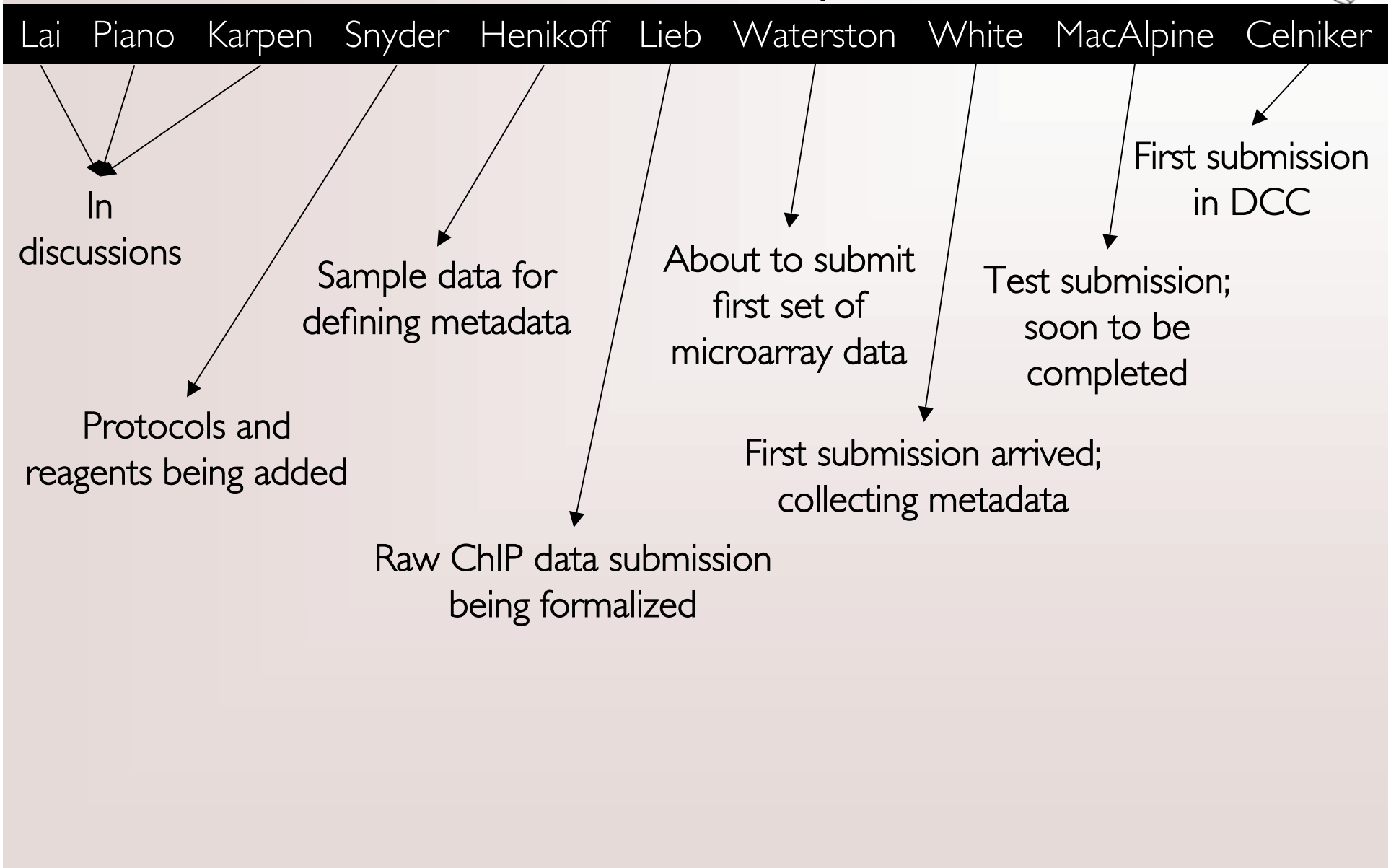
Why Consistency

- Comparison, CompaRison, Comparison
- Without consistency there can be no comparison





Status of data delivery



modENCODE Q & A

- Royal Palm Salon 4
- Personal question and answer sessions
- Ask for Richard, Nicole, Gos or EO





RIVERWALK GOLF CLUB
← Town and Country Guests
Have Preferred Tee Time Availability.
Dial Ext 1234 for information.
Parking Gate
Parking
Parking
GOLDEN PACIFIC DRIVE
ROYAL PALM DRIVE
ACCESSIBLE RESTROOMS
ACCESS ROUTE
ACCESSIBLE GUEST ROOM LOCATIONS ROYAL PALM TOWER 3500 ROOMS
ATM AUTOMATED TELLER MACHINE
ROYAL PALM TOWER
1000 - 1016 10th Floor 1017 - 1033
1900 - 1916 9th Floor 1917 - 1933
1800 - 1816 8th Floor 1817 - 1833
1700 - 1716 7th Floor 1717 - 1733
1600 - 1616 6th Floor 1617 - 1633
1500 - 1516 5th Floor 1517 - 1533
1400 - 1416 4th Floor 1417 - 1433
1300 - 1316 3rd Floor 1317 - 1333
1200 - 1216 2nd Floor 1217 - 1233
1100 - 1116 1st Floor 1117 - 1733
GOLDEN BALLROOM
GOLDEN PACIFIC BALLROOM
PACIFIC SALON ONE
PACIFIC SALON TWO
PACIFIC SALON THREE
PACIFIC SALON FOUR
PACIFIC SALON FIVE
PACIFIC SALON SIX
PACIFIC SALON SEVEN
PACIFIC BALLROOM
FOYER
LION FOUNTAIN COURT
FOOD COURT
Sunshine Deli
Terrace Cafe
WHIRLPOOL
POOL
GIFT SHOP
TIKI PAVILION
ROSE GARDEN
DOVER
STRATFORD
ATLAS BALLROOM
TOWN & COUNTRY
SAN DIEGO
GOLDEN WEST
CALIFORNIA
ATLAS FOYER
ROYAL PALM BALLROOM
ROSE PARK LANE
SUNSET
MEETING HOUSE
SUNRISE
ESQUIRE
TOWNE
DEVONSHIRE
TIFFANY ROOM
Trellises
Garden Grille
GRAND EXHIBIT HALL
TRELLISES COURTYARD
ROYAL PALM WAY
ROYAL PALM DRIVE
REGENCY LANE
REGENCY ROAD
FASHION VALLEY ROAD
Parking Structure
REGENCY TOWER
LE PAVILLON
LE SOMMET
WINDSOR ROSE
LE CHAMPLAIN
2800 - 2831
2700 - 2731
2600 - 2631
2500 - 2531
2400 - 2431
2300 - 2331
2200 - 2231
2100 - 2131
3580 3571
3540 3531
3541 - 3543
3500 - 3502
3544 - 3546
3503 - 3505
3530 3570
3517 3557
3506 3547
3516 3556
3450 3462
3400 3414
3749 - 3783
3735 - 3769
3756 - 3789
3742 - 3775
SHEFFIELD
HAMPTON
WINDSOR
GARDEN SALON ONE
GALLERIA ONE
GARDEN SALON TWO
GALLERIA TWO
CRESCENT
ASCOT
BRITTANY
CLARENDON
EATON
FAIRFIELD
3722 3733 3710
3711 3723 3700
Charlie's Trolley
3265 - 3268
3260 - 3263
3331 3383
3325 3375
3374
3364
3624
3661
3209 3210
3219
3150
3227 3237
3247 3257

www.modencode.org/cgi-bin/gbrowse/fly/