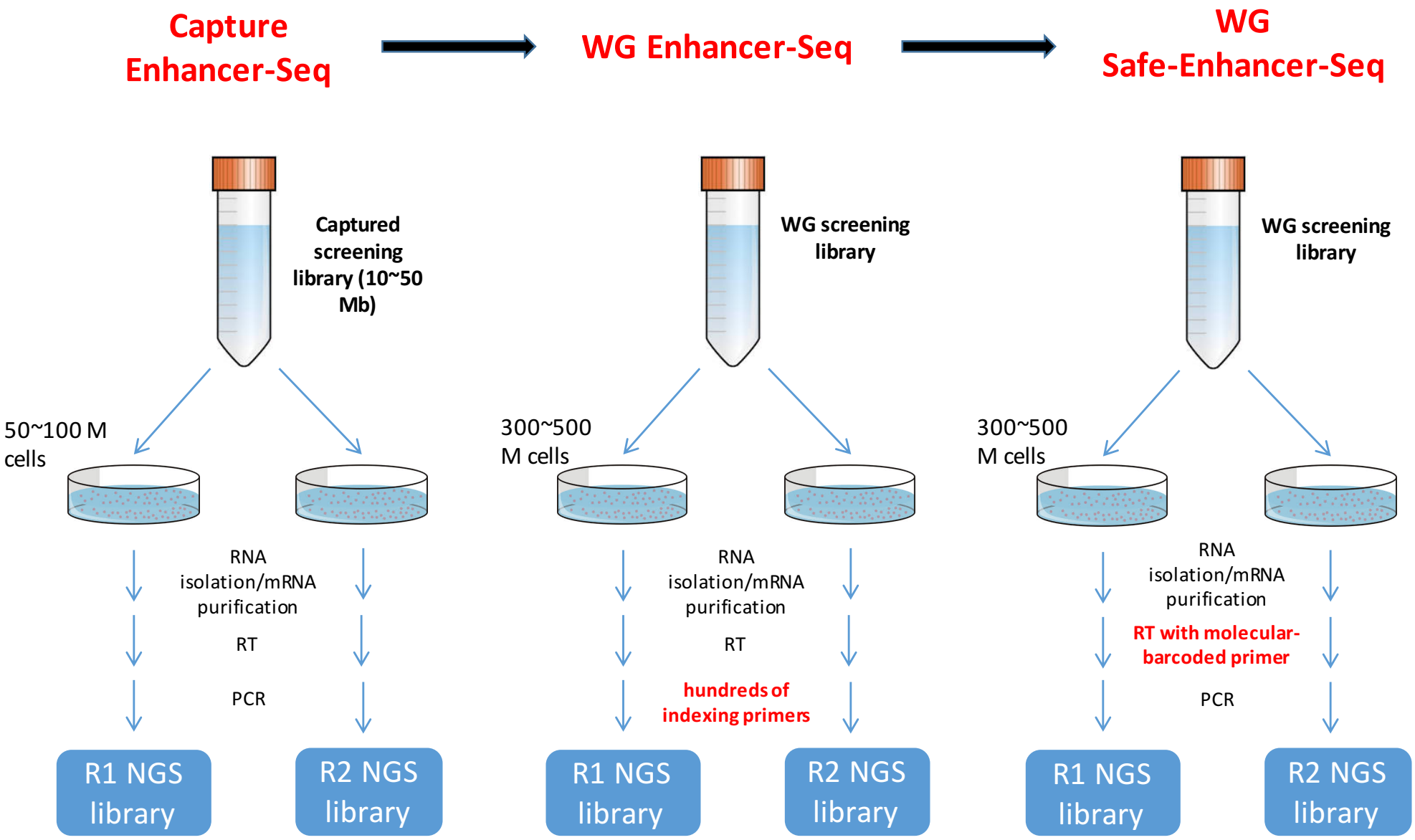
Enhancer-Seq

ENCODE Functional Characterization/Validation

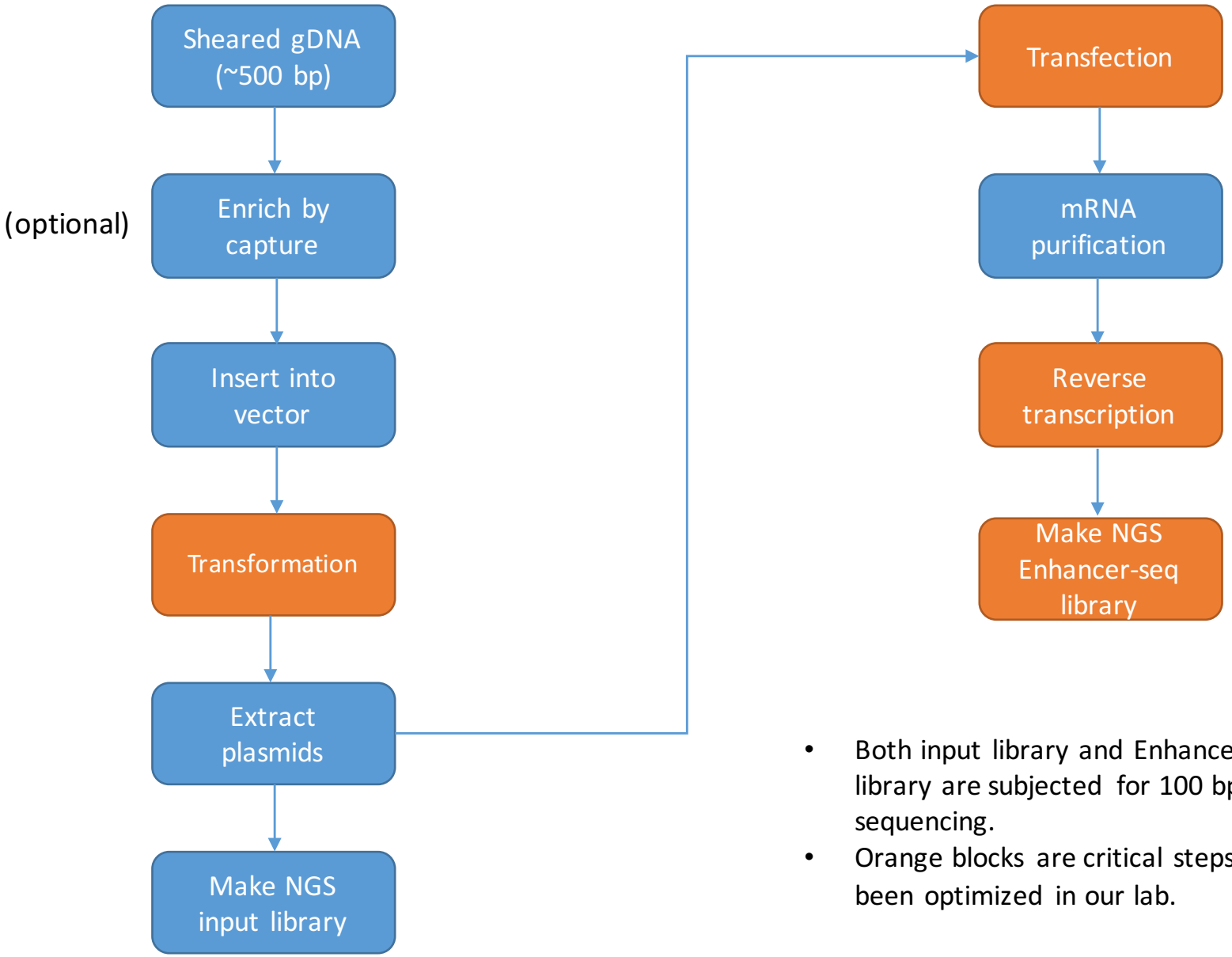
Kevin White Lab

11-20-2015

White Lab Enhancer-seq strategies

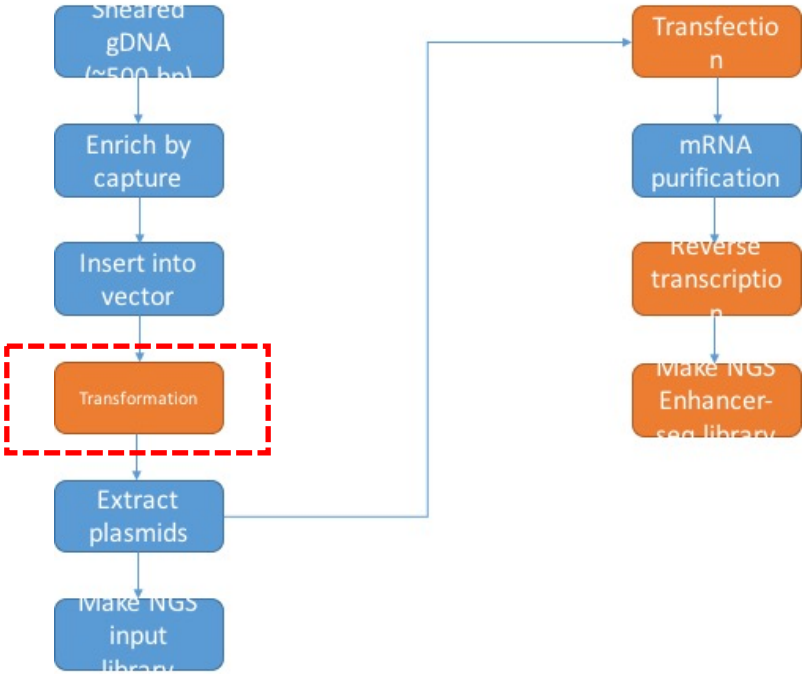


Experiment overview



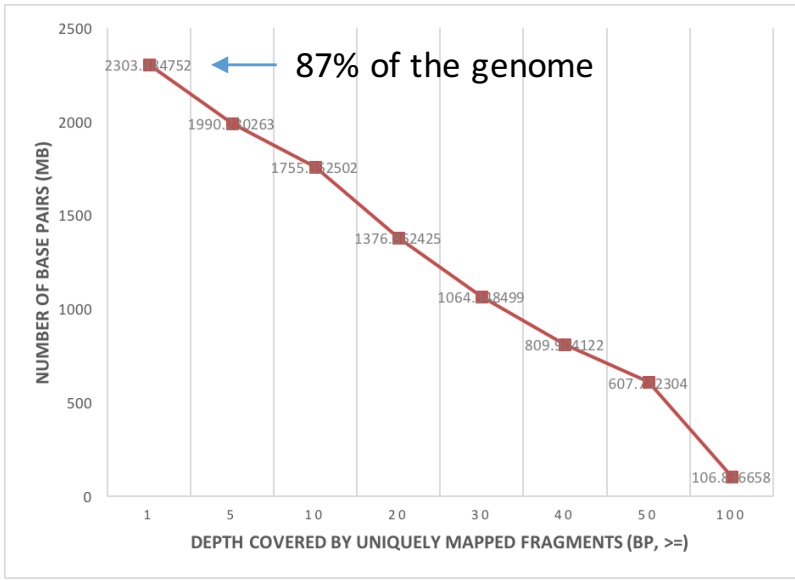
- Both input library and Enhancer-seq library are subjected for 100 bp PE sequencing.
- Orange blocks are critical steps that have been optimized in our lab.

Major Improvements - 1

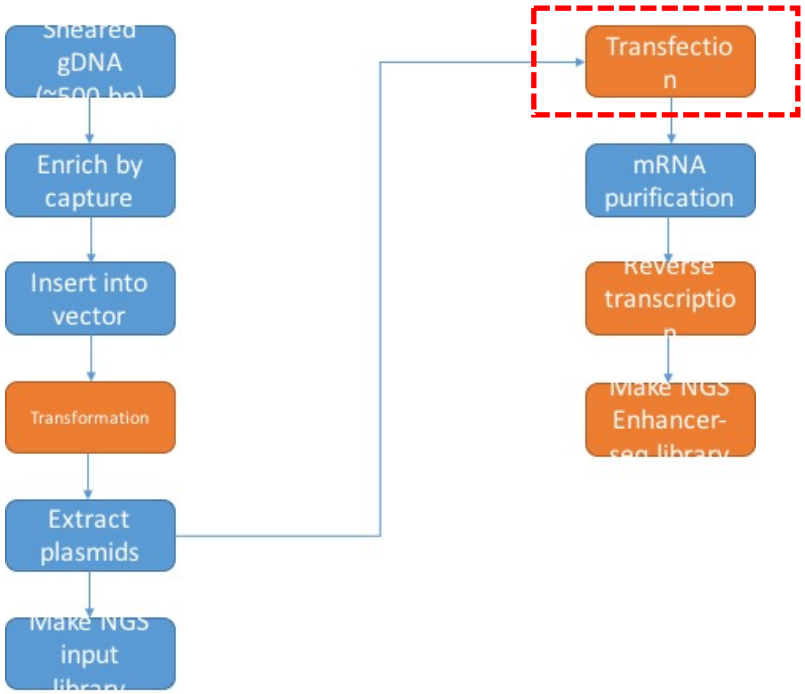


Transformation condition has been largely optimized, so that the current efficiency allows us to make high complex plasmids pool for whole-genome enhancer screening.

- >= 15 million clones per Gibson Assembly
  - Given transformation from 48 GA reactions, the expected genome coverage is between 0.24x to 120x



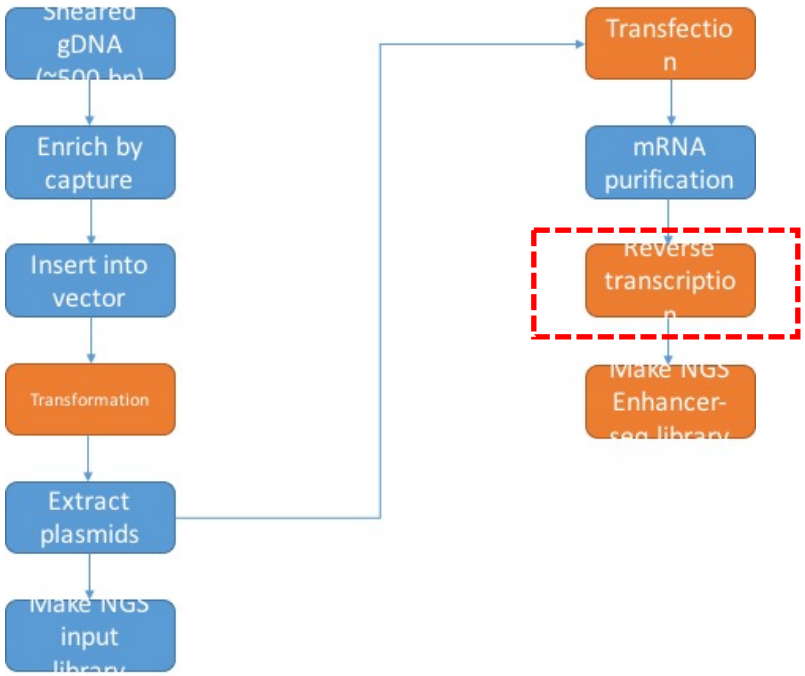
Major Improvements - 2



Transfection conditions have been optimized and scaled up to allow transfect 200~500 million cells within two hours.

- LCL (GM12878 etc)
- MCF7
- T47D
- LNCaP
- SNU16
- OCUM1
- HEK293
- K562 (under testing)

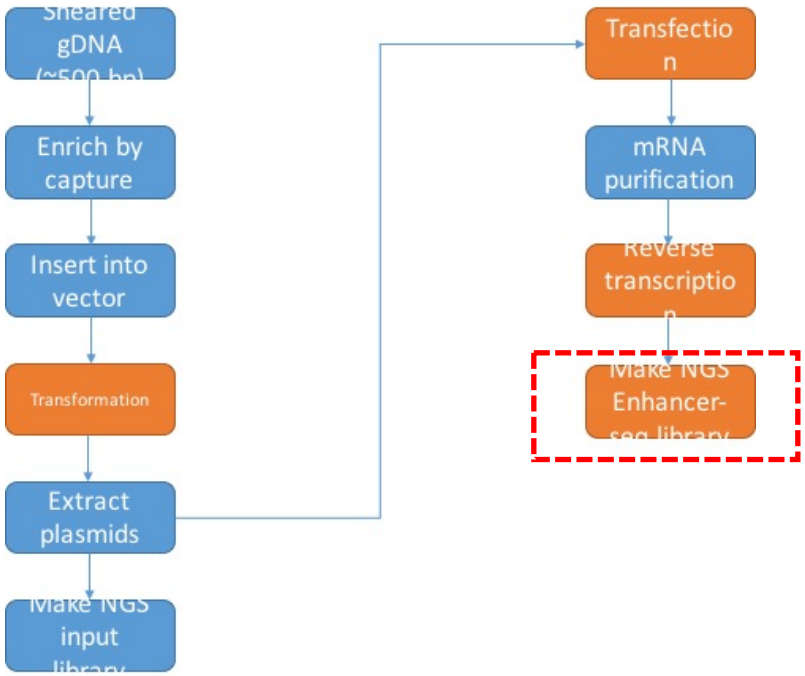
Major Improvements - 3



Random molecular barcodes were added to distinguish biological copies of plasmid-transcribed-mRNA from PCR duplicates.

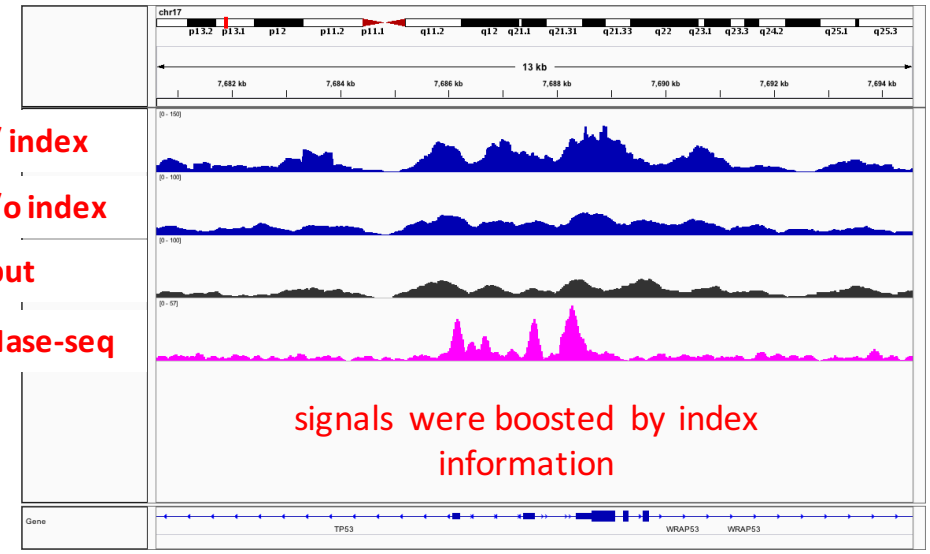
- Safe-Enhancer-Seq (has some preliminary data in whole-genome LCL experiments)

Major Improvements - 4

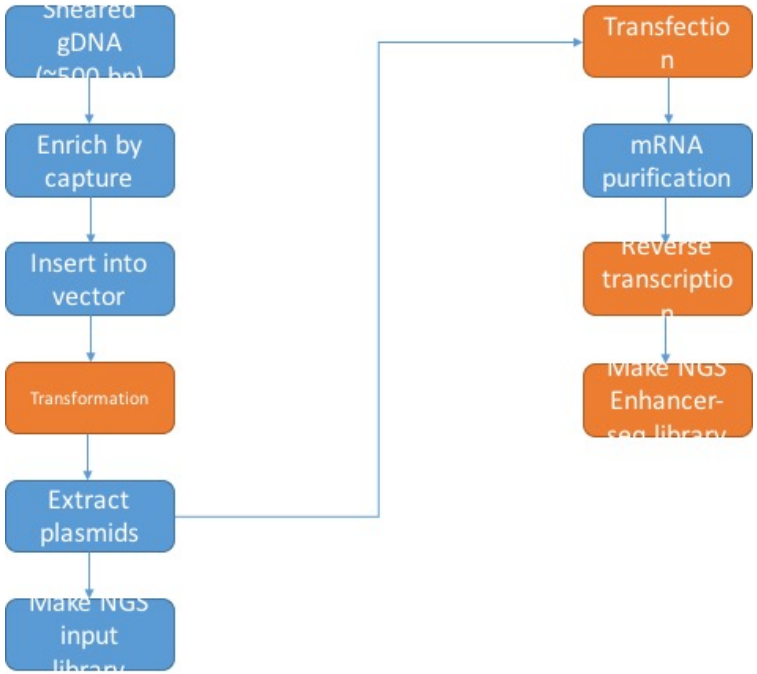


Hundreds of indexing-primers are used to distinguish biological copies of plasmid-transcribed-mRNA from PCR duplicates (alternative of Safe-Enhancer-Seq).

- Full data sets are available in multiple LCL (whole-genome).
- More data coming out in SNU16 and OCUM1 (capture).



Other Improvements



Multiple vectors (with different promoters and reporter genes) were tested.

- SCP1
- miniCMV
- minP

- Luciferase
- GFP

Multiple kits/reagents were tested in most steps to balance:

- Efficiency
- Scale
- Cost

Establish check-points in different steps to ensure data quality.

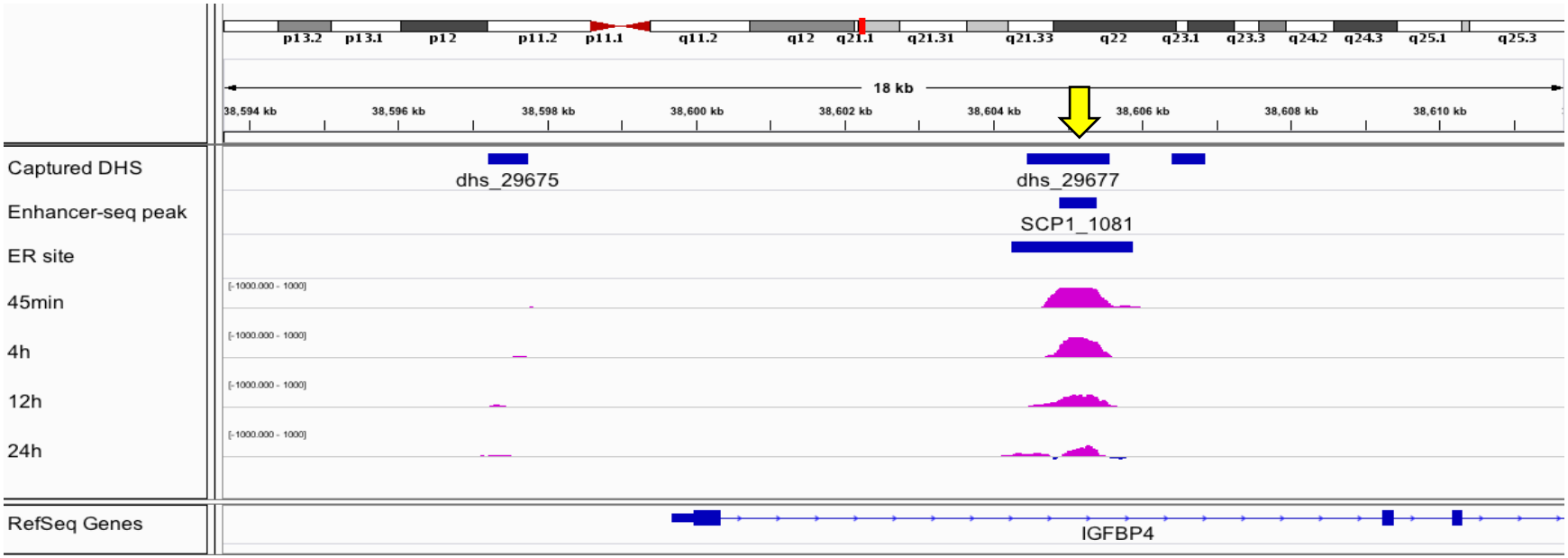
Ongoing projects

- **Nuclear receptor binding sites capture sets (50 Mb, 100,000 regions)**
  - LNCaP (vehicle and DHT-treated)
- **E2-reponsive DHS capture sets (12 Mb)**
  - MCF7 (vehicle and E2-treated, time-course)
- **Gastric Cancer capture sets (40 Mb) (collaboration)**
  - SNU16
  - OCUM1
- **Evolved/diverged DHS capture sets (15 Mb)**
  - LCL
- **Whole-genome**
  - LCL
  - LNCaP (vehicle and DHT-treated)
  - K562
- **Whole-genome Safe-Enhancer-Seq**
  - LCL

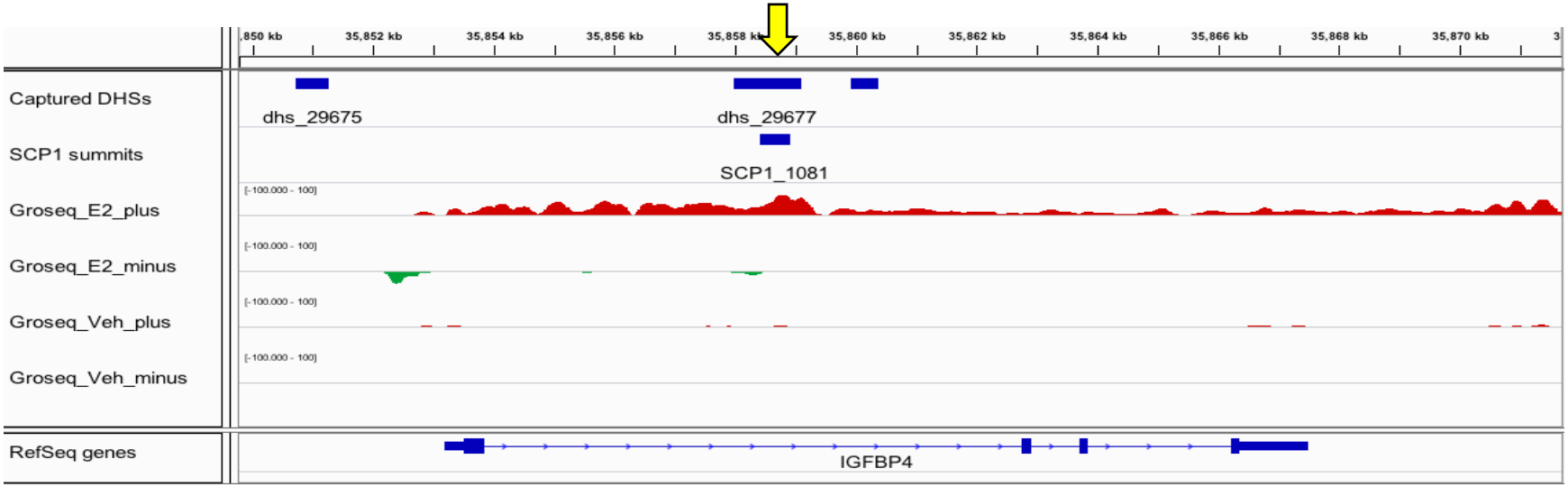
Examples of E2-induced Enhancers

IGFBP4

Enhancer-seq



GRO-seq 1hr



Timeline for K562

- **Week 1**
  - **Prepare WG input library**
  - **Send for sequencing (verify quality)**
- **Week 2-3**
  - **Testing transfection condition**
  - **Grow cell population to large number**
- **Week 4**
  - **Large-scale transfection**
  - **Make WG Enhancer-seq library and send for sequencing**
- **Wee 5-8**
  - **Deep sequencing**
  - **Data analysis (with Mark Gerstein's Group)**

We're expecting to get the preliminary results of K562 enhancer actvity by the end of December.