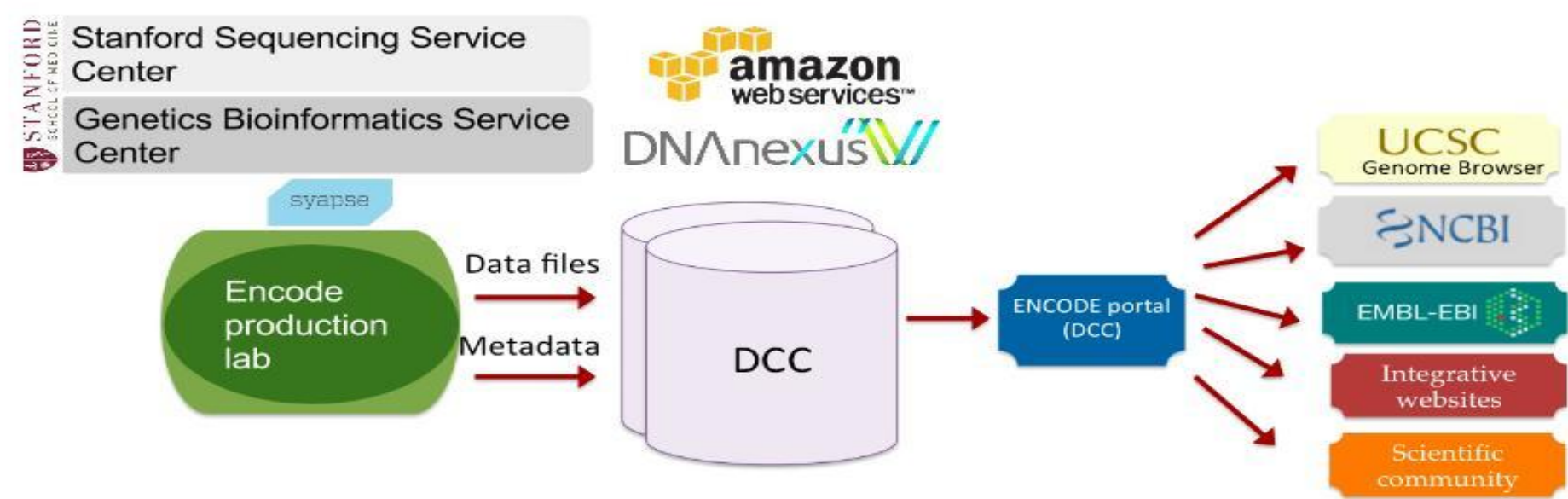


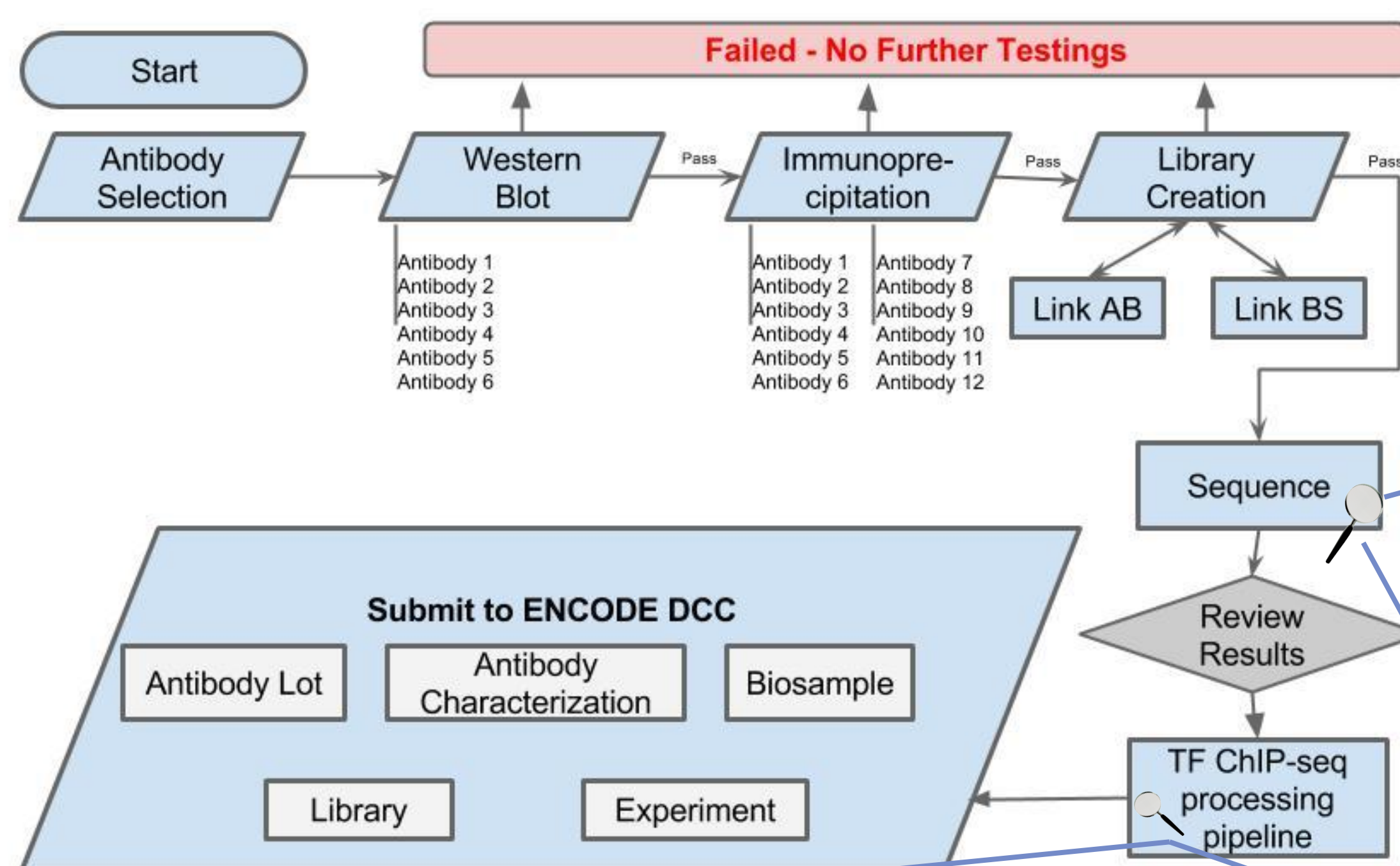
# Informatics for Encyclopedia of DNA Elements (ENCODE3)

DENIS N. SALINS, NATHANIEL K. WATSON, ESTHER CHAN, SOMALEE DATTA, TRUPTI KWALI  
Stanford University

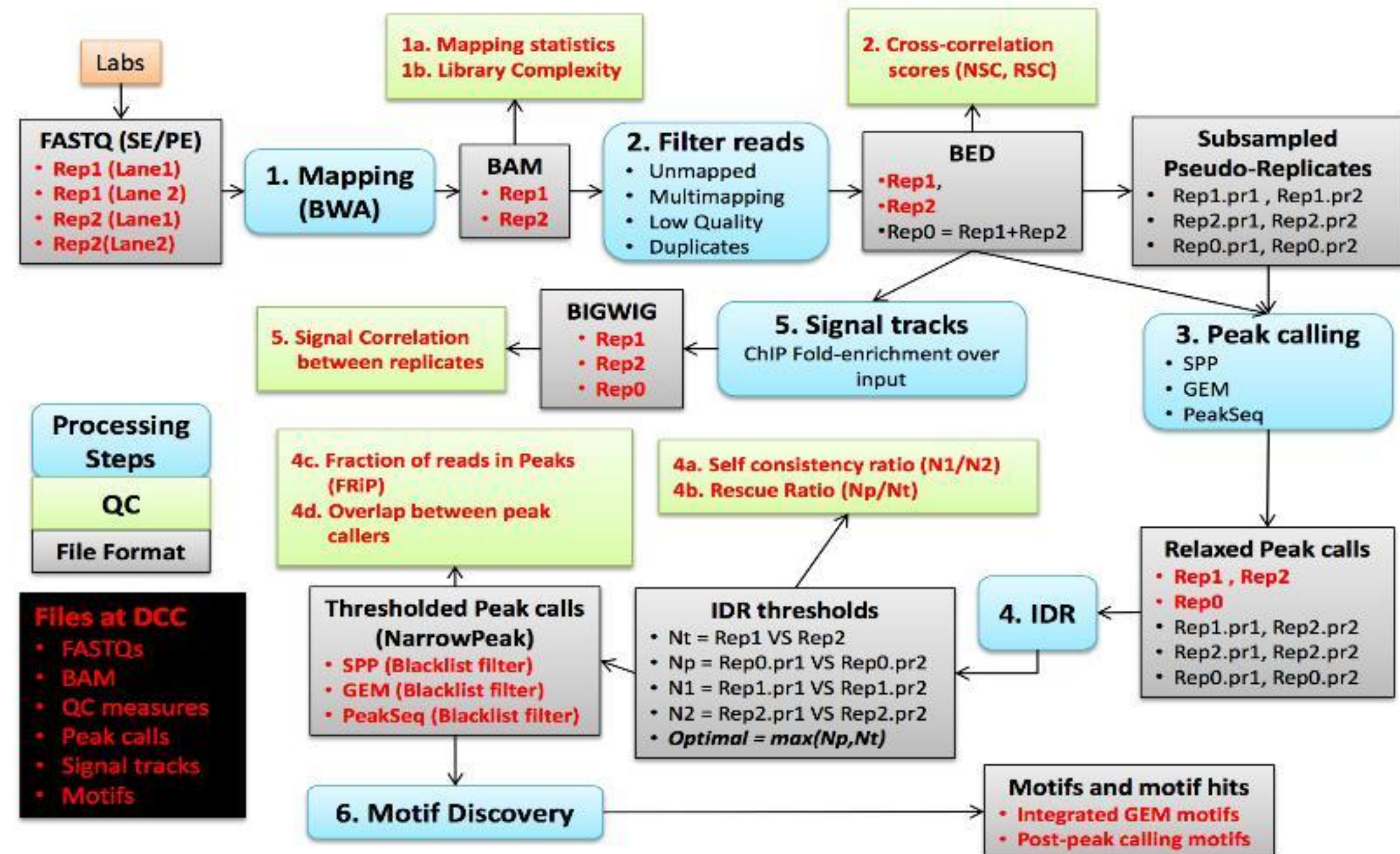


Encyclopedia of DNA Elements (ENCODE) project is a massively large scale collaborative research effort, to create a comprehensive public catalog of functional elements in the human genome. All data generated from production labs are submitted and presented by the Data Coordination Center (DCC). The DCC handles the data validation, tracking, storage, distribution, and visualization to the research community. Stanford is amongst the largest contributors to the overall project with submissions of tested Antibodies, Transcription Factors, and Biosamples to produce over 10,000 of data and meta data points.

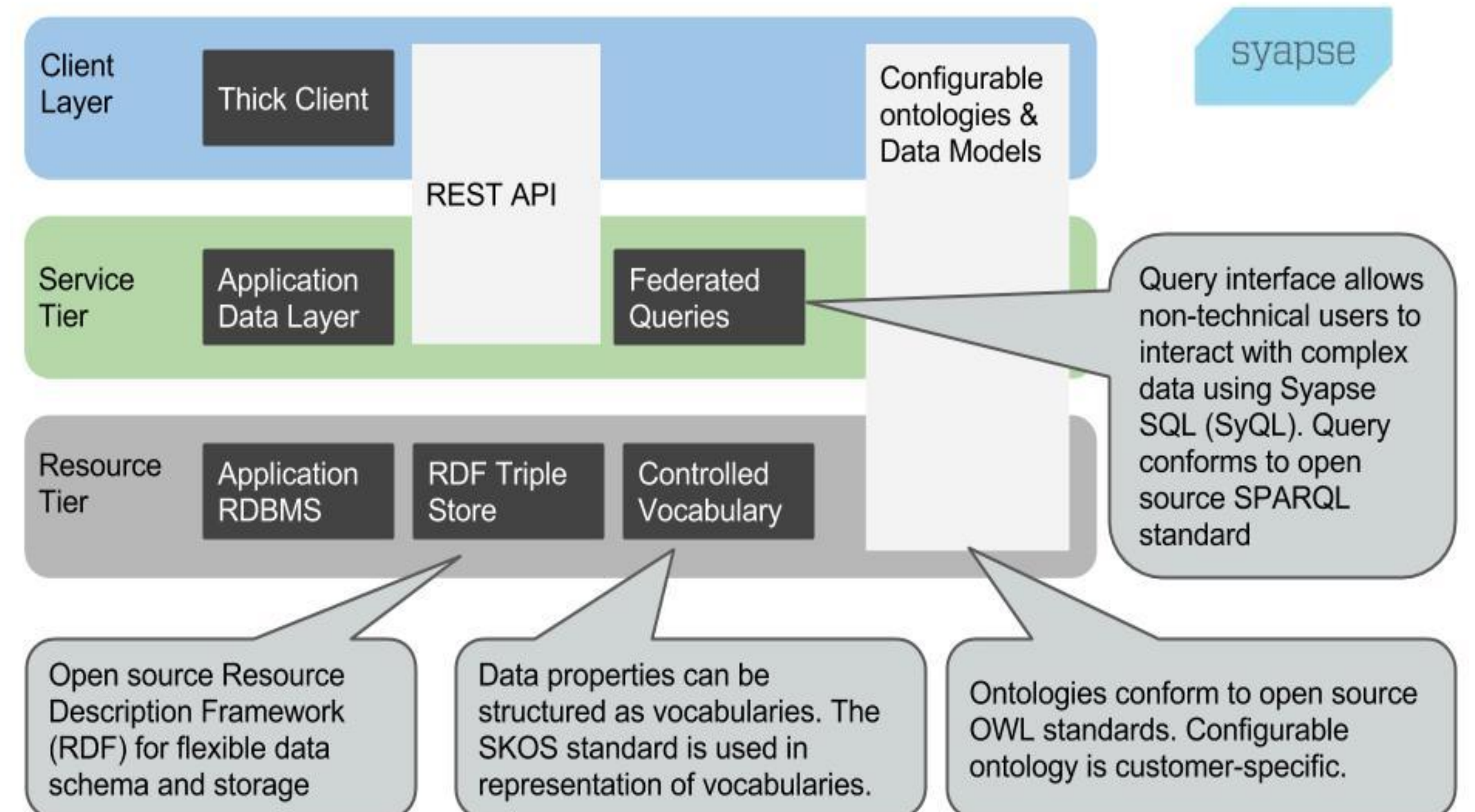
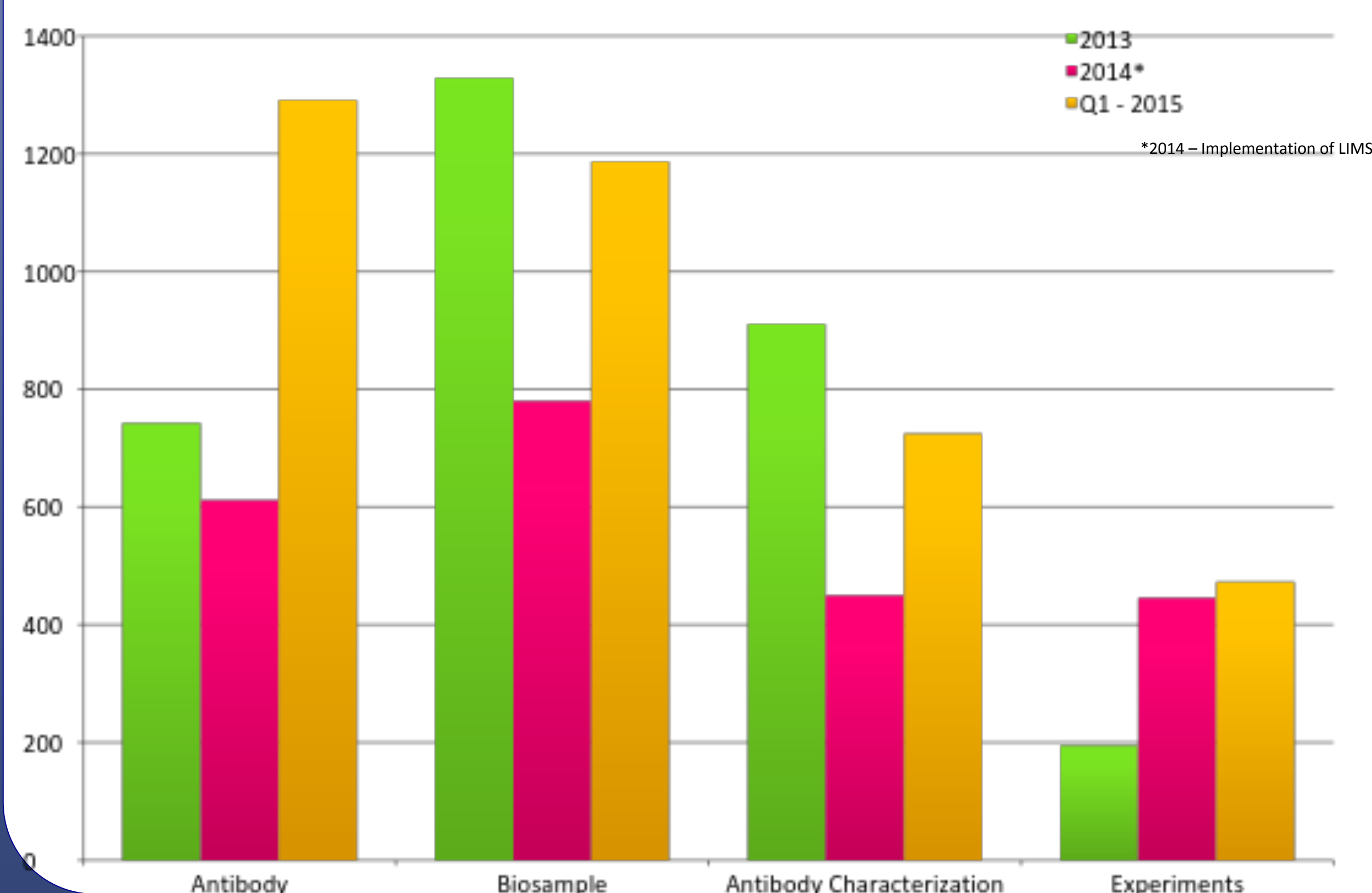
The data generated to achieve these datasets is also immensely important to understand the process that was used to test and arrive at the published results. Typically, testing 100 antibodies will yield a 5-15% (Testing over 200,000) success factor of the transcription factor passing all the desired tests and experiments.



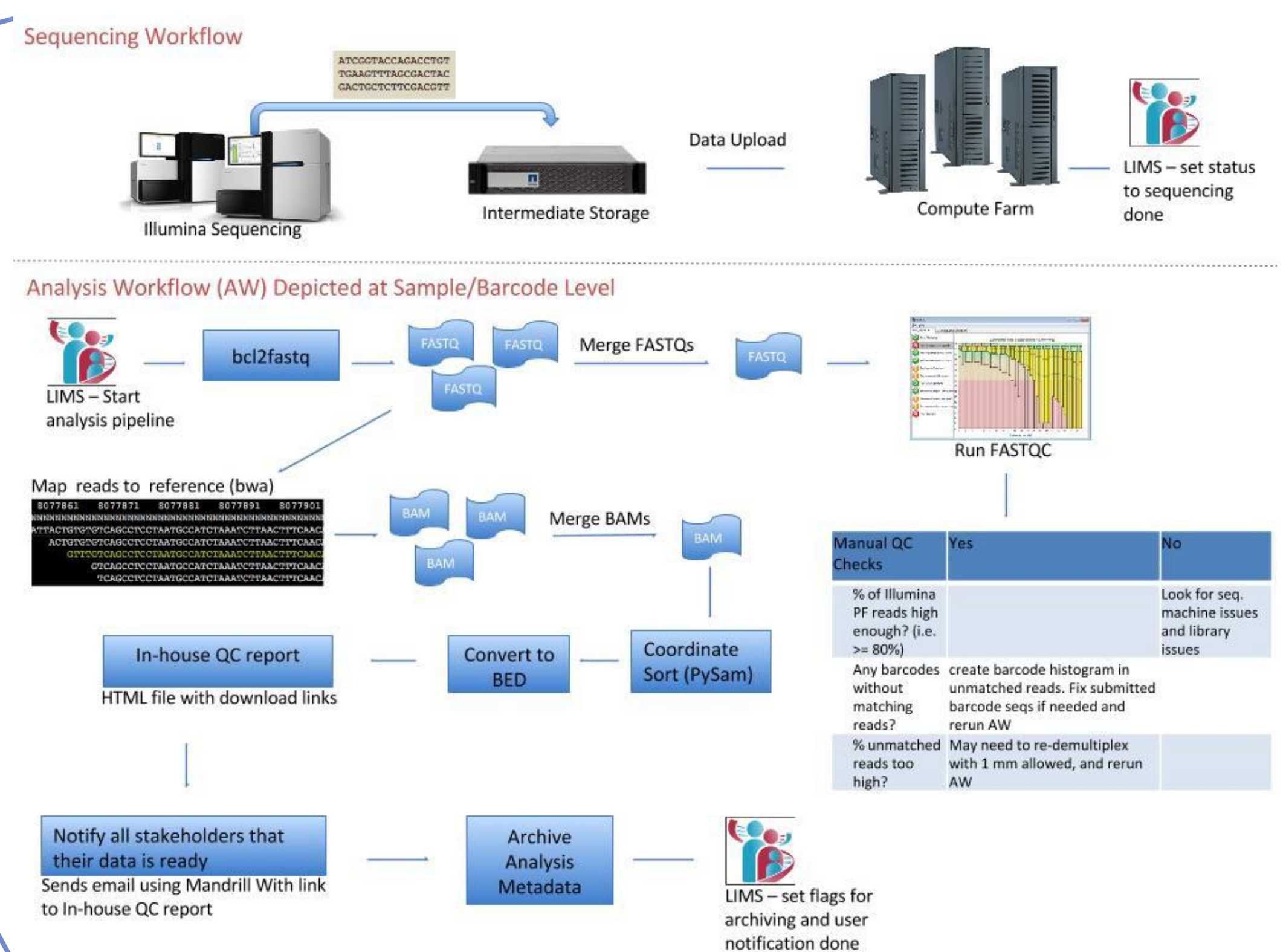
## TF ChIP-seq processing pipeline



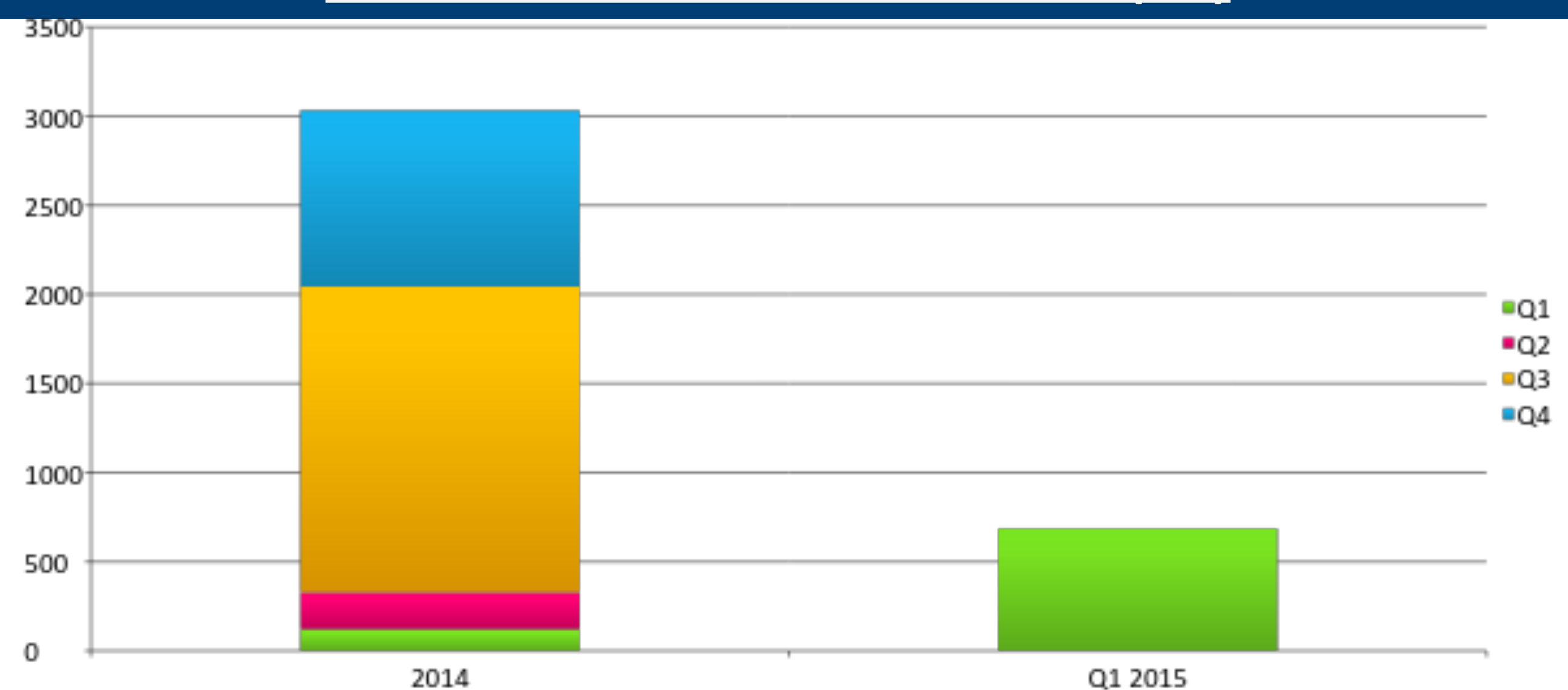
## Stanford – ENCODE Submissions



## Stanford's – Sequencing Workflow



## Stanford's – ENCODE 3 Data Scale (GB)



## Stanford's ENCODE Contribution

