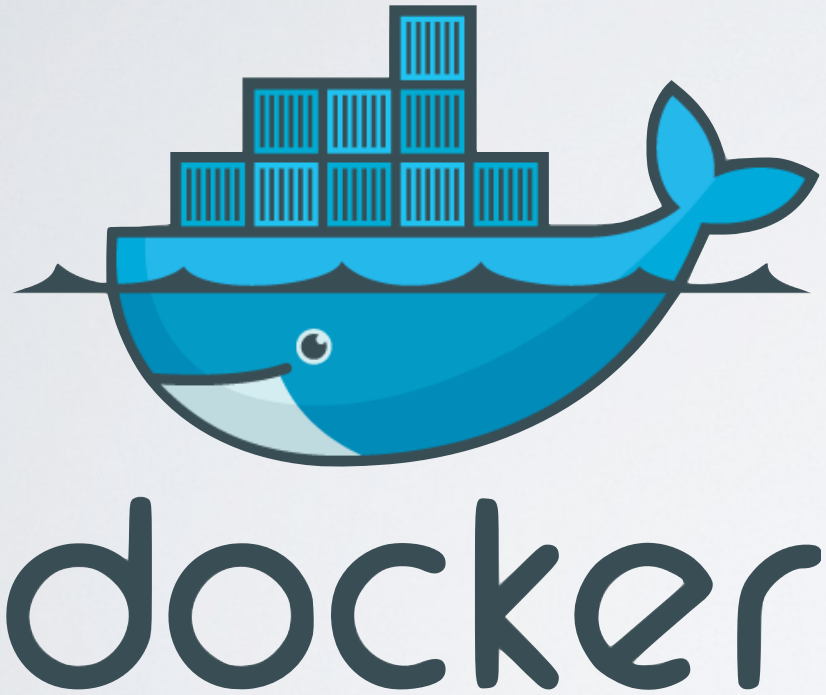
JC

docker + BaseSpace

RK

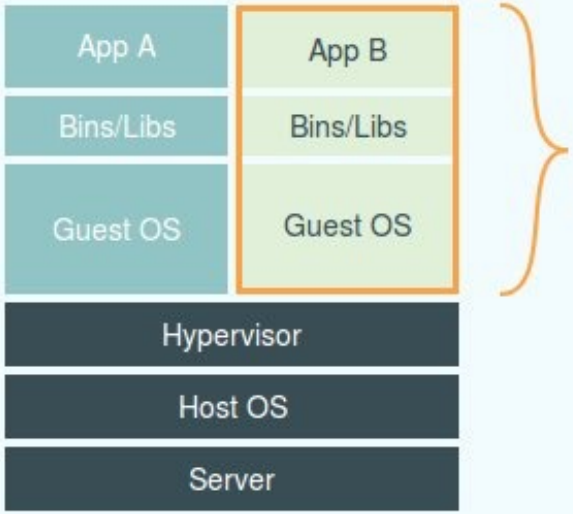
2015 - 01 - 28

what is docker?



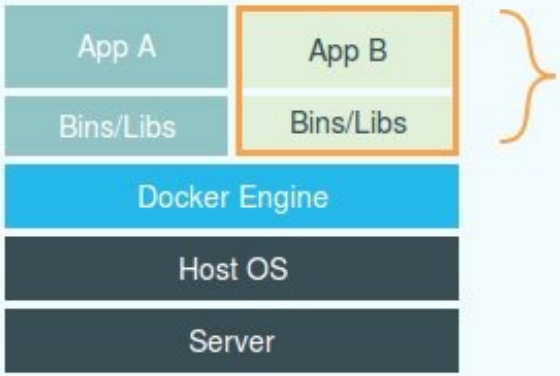
- Docker uses LinuX Containers (LXC), which run in the same operating system as its host
- This allows it to share a lot of the host operating system resources

VE vs. VM



Virtual Machines

Each virtualized application includes not only the application - which may be only 10s of MB - and the necessary binaries and libraries, but also an entire guest operating system - which may weigh 10s of GB.



Docker

The Docker Engine container comprises just the application and its dependencies. It runs as an isolated process in userspace on the host operating system, sharing the kernel with other containers. Thus, it enjoys the resource isolation and allocation benefits of VMs but is much more portable and efficient.

getting started!

- docker supported natively on Red Hat Enterprise 7
- for Windows/Mac, use **boot2docker**: http://boot2docker.io

using docker

- two modes for creating a docker image:

  1) interactively using **`docker run`** and **`docker commit`**

  2) statically using **`docker build`**

- I prefer the static build for **reproducibility** and **clarity**

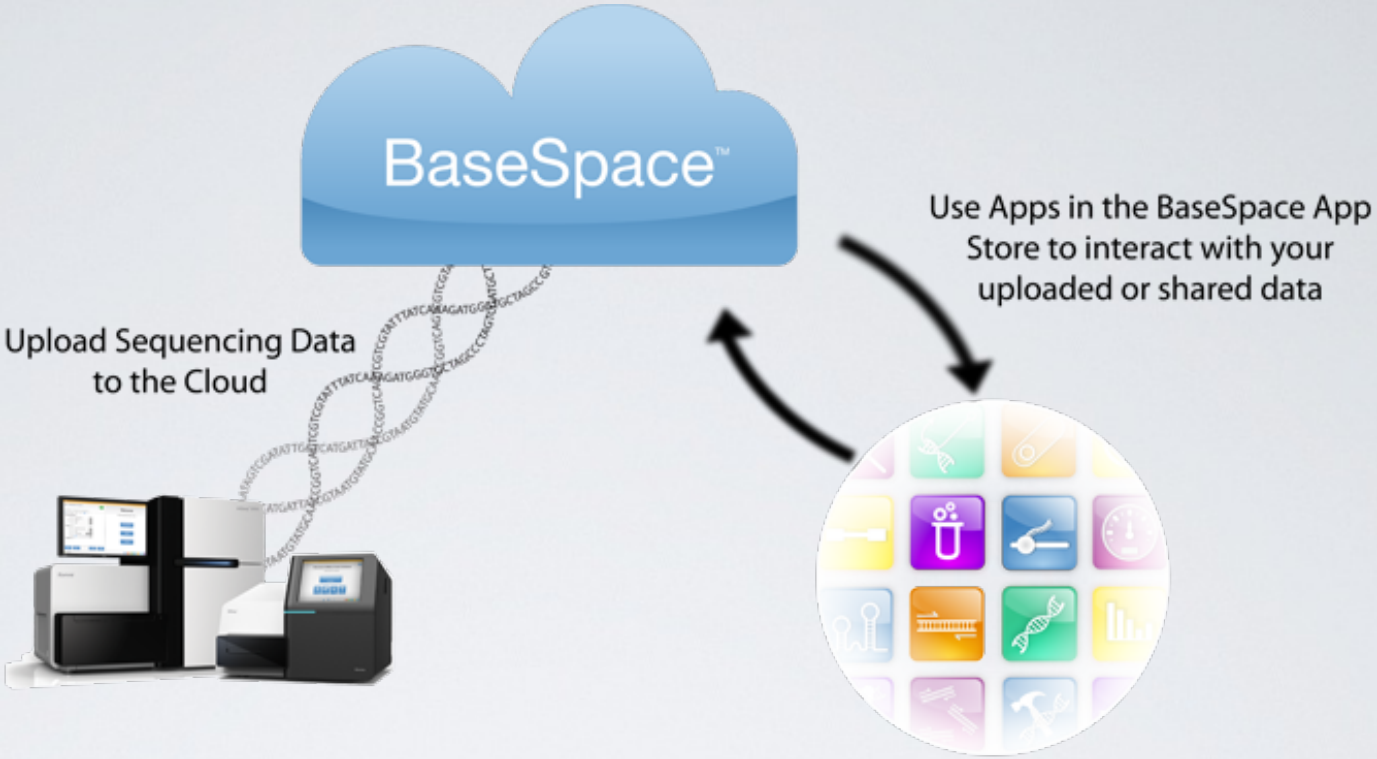
getting started!

```
$ sudo docker search ubuntu

$ sudo docker pull ubuntu

$ sudo docker run -i -t ubuntu /bin/bash
```

- docker has a nice web-based interactive tutorial: https://www.docker.com/tryit

basespace app

simple app, slightly more complicated docker build…



- Illumina BaseSpace
- nice GUI for genomic apps, runs in AWS
- apps packaged and deployed using docker

the central dogma is cell-type specific

dynamic

RNA turnover

protein turnover
PTMs

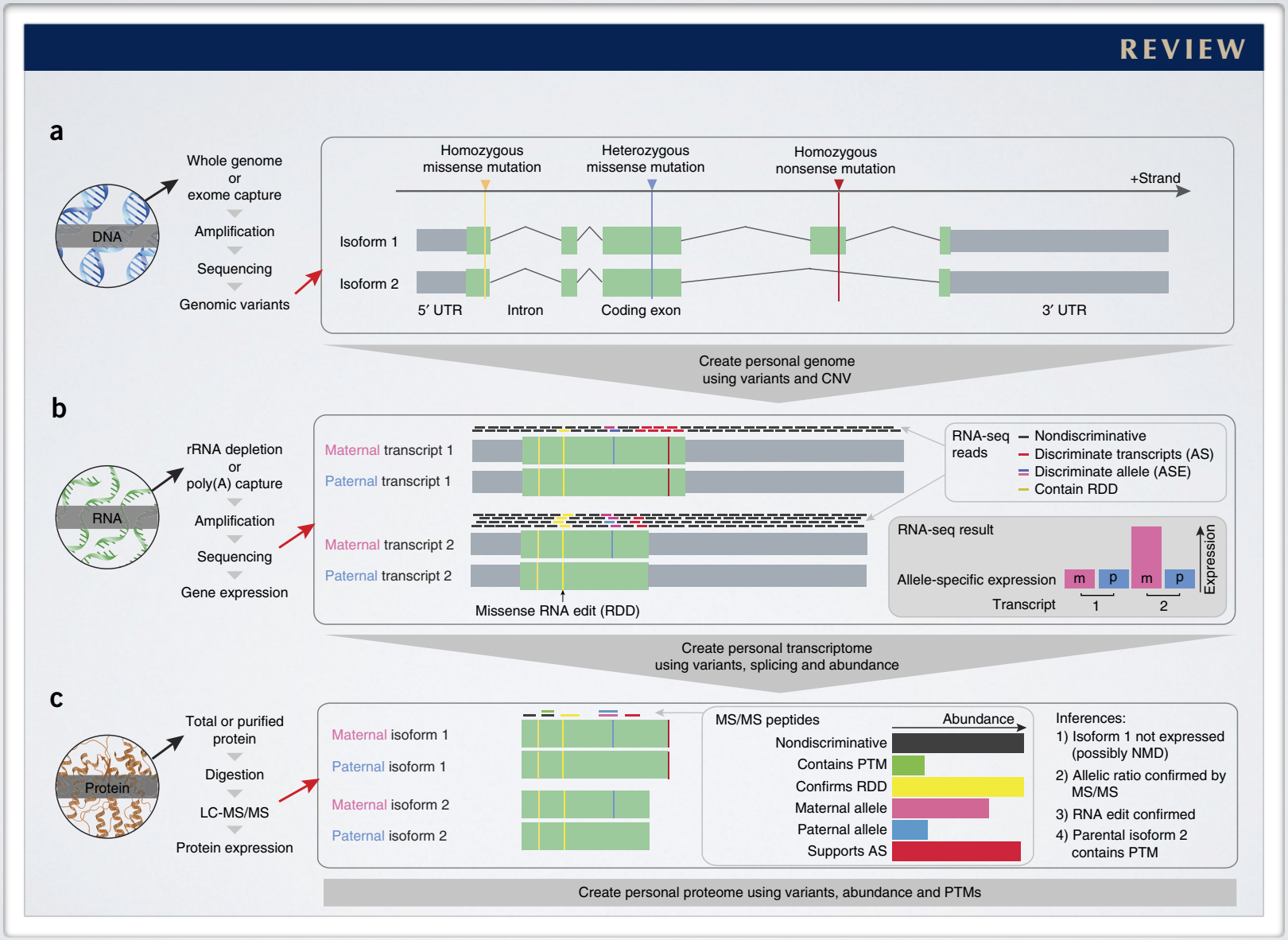
structural variants / SNPs
promoter usage / splicing
epigenetics / imprinting

RNA abundance
RNA editing
miRNA repression
NMD

protein abundance

static

high-throughput profiling of the CNS: the case for deeper *omic integration



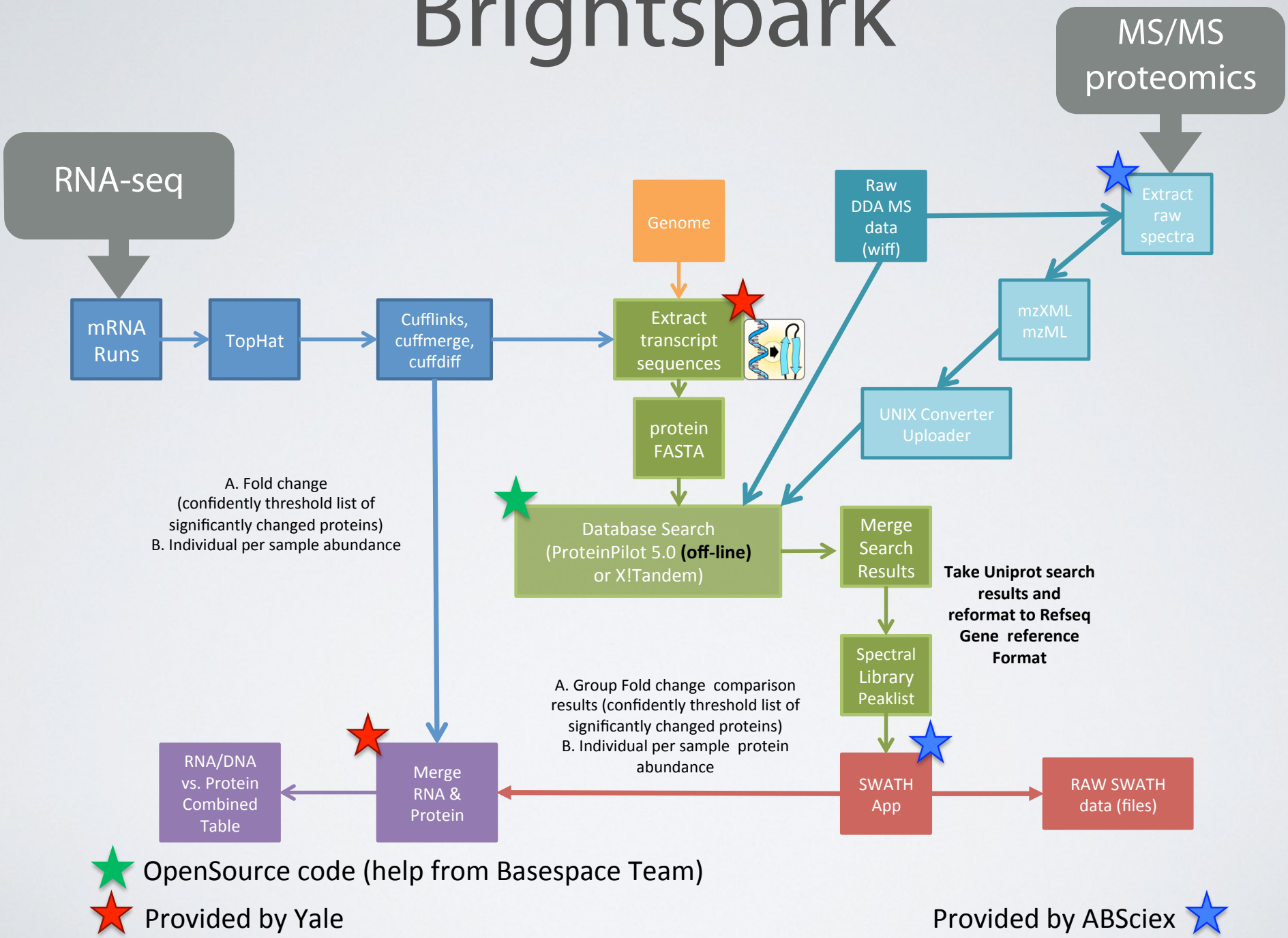
my app

- '**RNA-Seq Translator**'
- developed as part of the much larger '**Brightspark**' collaboration between Illumina, ABSciex, and Yale
- aim to more meaningfully **integrate RNA & protein analysis**

cufflinks (cuffdiff) output (.gtf)

protein fasta of cufflinks transcripts

Brightspark



app @ basespace