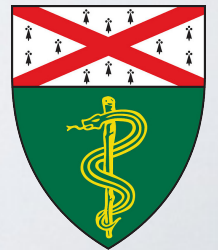


JC

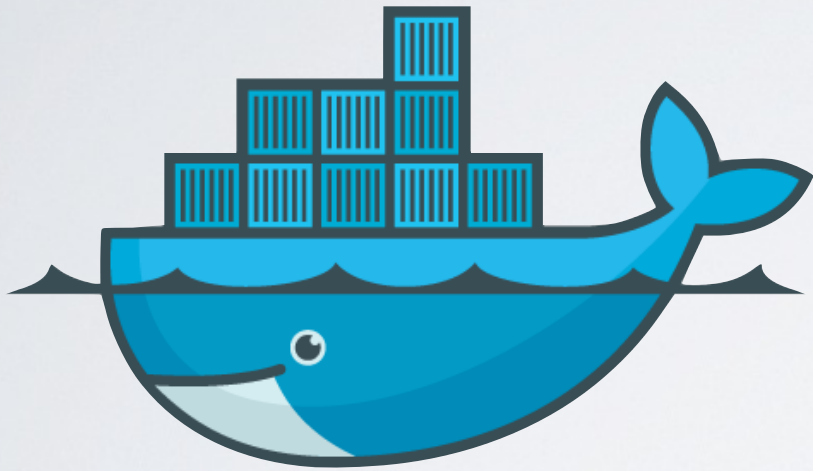
docker + BaseSpace

RK

2015 - 01 - 28



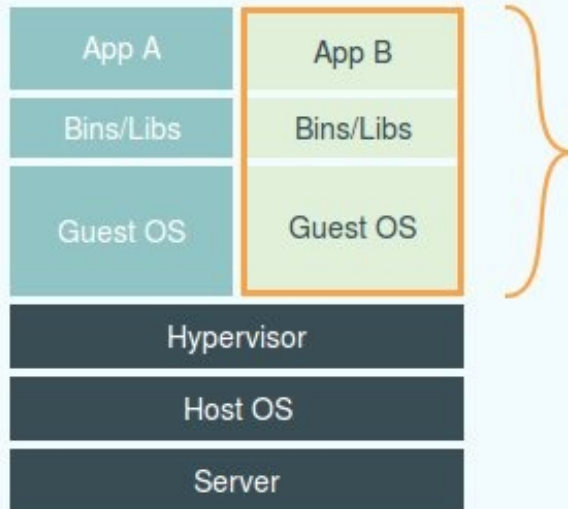
what is docker?



docker

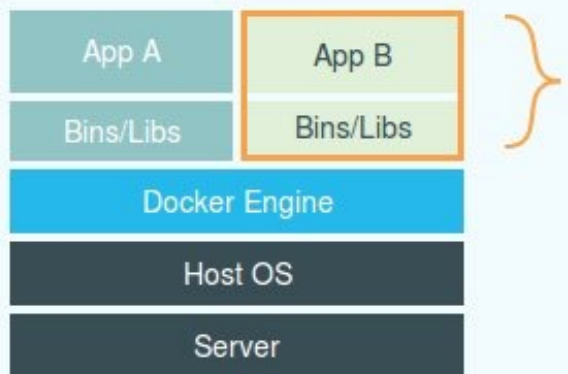
- Docker uses Linux Containers (LXC), which run in the same operating system as its host
- This allows it to share a lot of the host operating system resources

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
```

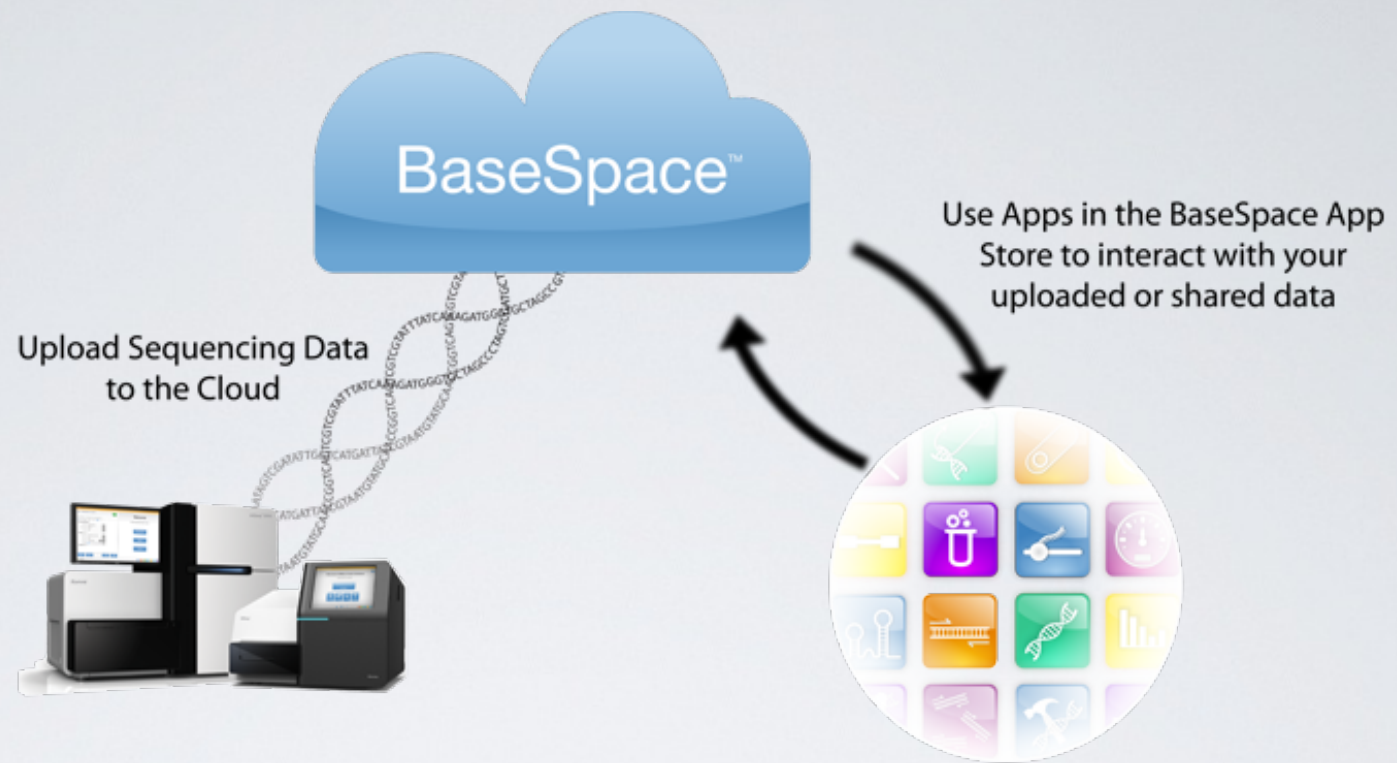
```
$ sudo docker run -i -t ubuntu /bin/bash
```

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



basespace app

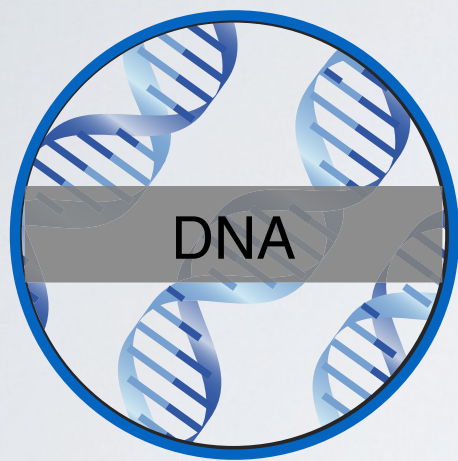
simple app, slightly more complicated docker build...



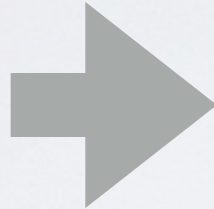
- Illumina BaseSpace
- nice GUI for genomic apps, runs in AWS
- apps packaged and deployed using docker

the central dogma is cell-type specific

dynamic



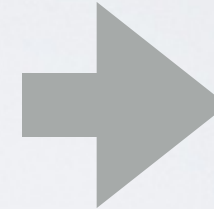
rates of
transcription




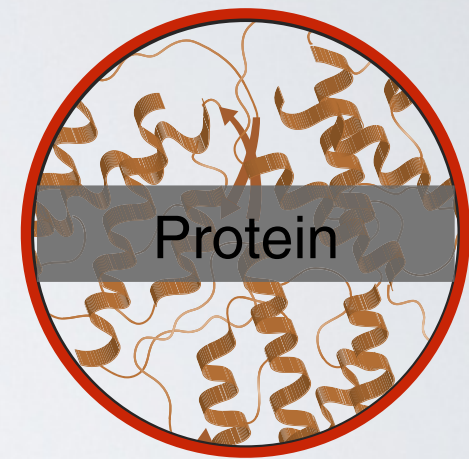
RNA turnover 



rates of
translation



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



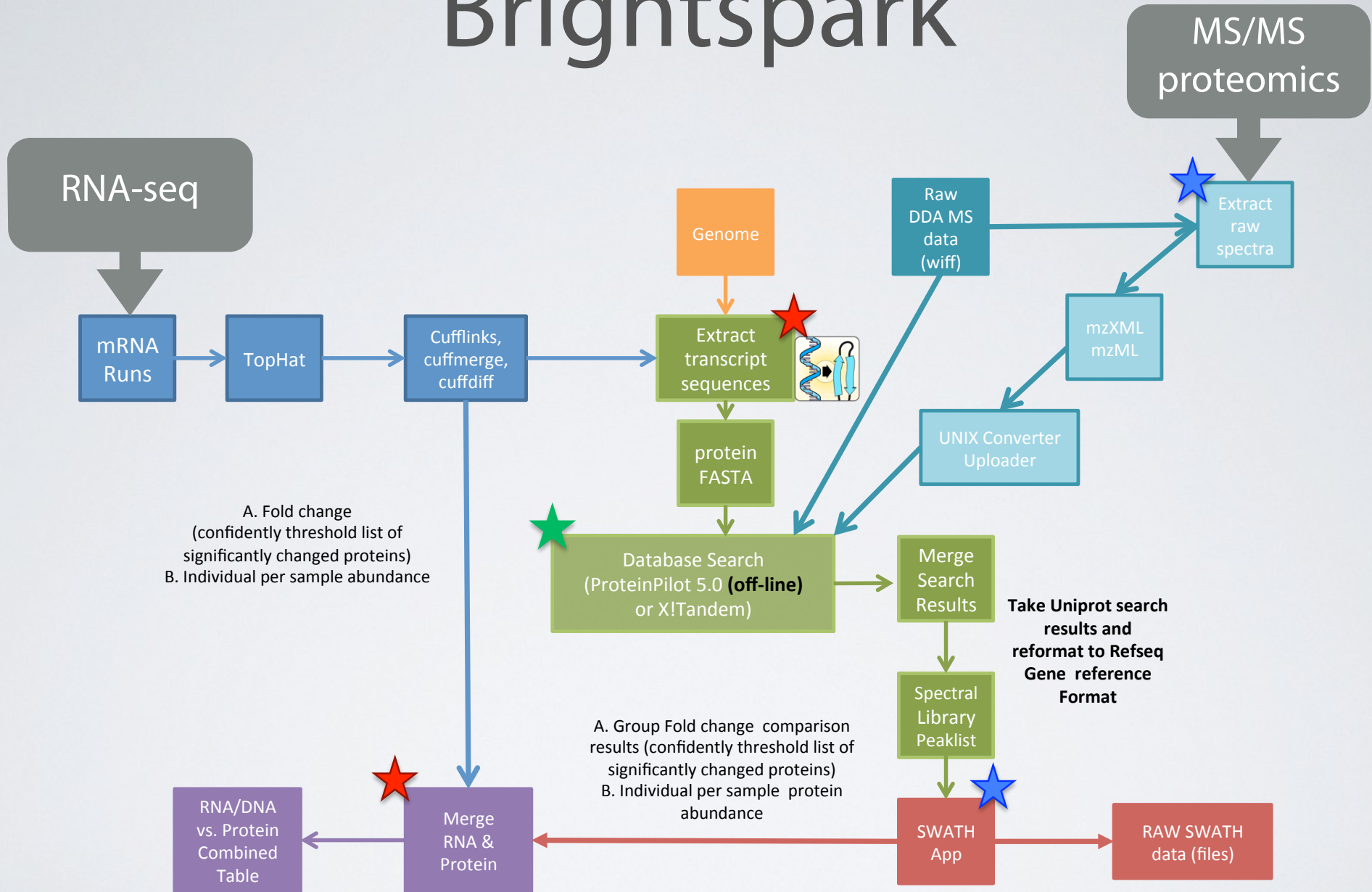
- 'RNA-Seq Translator'
- developed as part of the much larger 'Brightspark' collaboration between [Illumina](#), [ABSciex](#), and [Yale](#)
- aim to more meaningfully **integrate RNA & protein analysis**

cufflinks (cuffdiff)
output (.gtf)



protein fasta of
cufflinks transcripts

Brightspark



A. Fold change
(confidently threshold list of significantly changed proteins)
B. Individual per sample abundance

A. Group Fold change comparison results (confidently threshold list of significantly changed proteins)
B. Individual per sample protein abundance

★ OpenSource code (help from Basespace Team)

★ Provided by Yale

Provided by ABSciex ★

app @ basespace

The screenshot shows a web browser window with the URL `https://basespace.illumina.com/appsession/start?clientId=b95fa29314b242d8908...`. The page header includes the BaseSpace logo and the Illumina logo. The main content area features the RNA-Seq Translator logo and the text "Yale University". Below this, there are three sections: "Analysis Name:" with a text input field containing "RNA-Seq Translator 01/26/2015 2:28:44"; "App Result:" with a dropdown menu showing "Cufflinks-Report"; and "Save Results To:" with a dropdown menu showing "RNA-seq -> Proteomics TEST". At the bottom right, there is a grey box with the text "This app is free." and a blue "Continue" button with a right-pointing arrow. A vertical "contact us" button is located on the right side of the page.