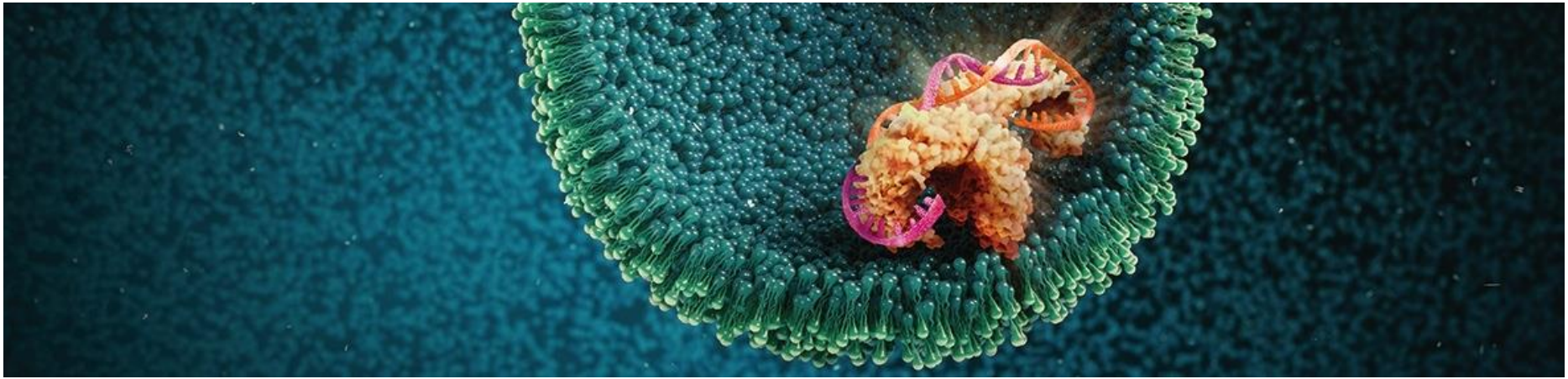
In Silico Biology at AstraZeneca

Master's Programme Presentation 2020

Sue Monkley

***Bioinformatics and Data Science**, Translational Science and Experimental Medicine | Research and Early Development | Early Respiratory & Immunology | BioPharmaceuticals R&D , AstraZeneca, Gothenburg, Sweden*

11 Dec 2020



At AZ we use cutting edge technology and tools to discover disease treatments for patients

We work in a highly collaborative environment

We interact with experts in chemistry, pharmacology and disease biology internally

We collaborate with external academic and clinical experts

We aim to make the best possible drugs which requires the highest quality science

Today we will focus on highlighting the way we work

- the role we, as in silico biologists, play in the drug discovery process
- particular challenges treating respiratory diseases
- specific examples
- future directions

AstraZeneca

‘by leading in science, we will transform the lives of patients around the world”

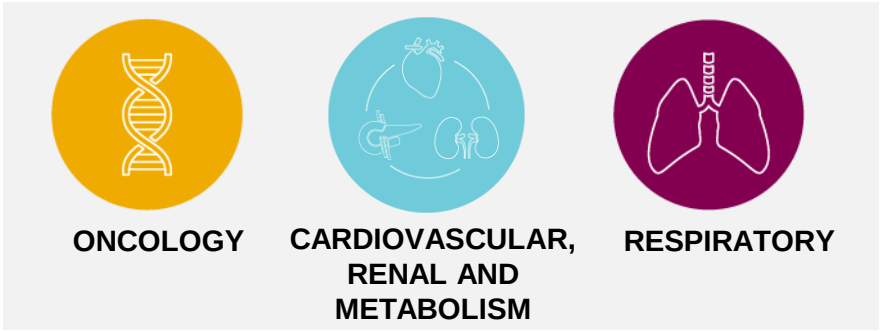
Pascal Soriot, CEO

61 100 employees around the world,

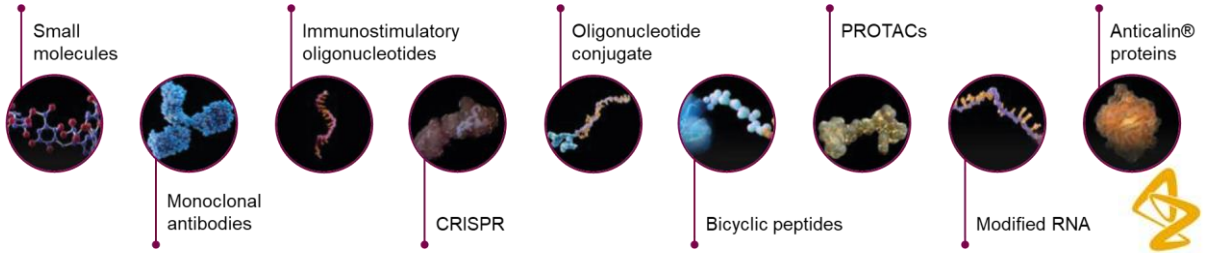
6 200 in Sweden

- **$5.9 bn** annually invested in Research & Development (across 5 countries)
- More than **600** collaborators and partnerships globally
- **625** publications (**44** high impact)- 2017 (IMED)
- **150** projects in our pipeline

MAIN THERAPY AREAS



A broad modality toolbox:



Gothenburg is one of the three strategic R&D centres within AstraZeneca

2 500 employees, **2** therapy areas

- 30 professors
- 600 with PhD
- 49 nationalities

100 collaboration projects (worldwide)

CARDIOVASCULAR, RENAL AND METABOLISM

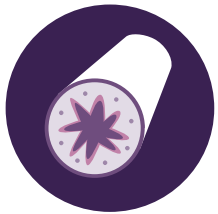
- Heart Failure
- Diabetes
- Chronic Kidney Disease

RESPIRATORY

- Asthma
- Chronic Obstructive Pulmonary Disease (COPD)
- Idiopathic Pulmonary Fibrosis (IPF)
- Chronic Cough

Early RIA: Our aim is to modify disease in areas of unmet need

We focus on four diseases with remaining unmet medical need and increasing prevalence



Asthma

~339M people

1 in 10 children at risk of developing asthma

COPD

~384M adults

50% dead 3½ years after 1st hospitalisation

IPF

Up to 100K diagnosed in US

80% of patients die within 5 years of diagnosis

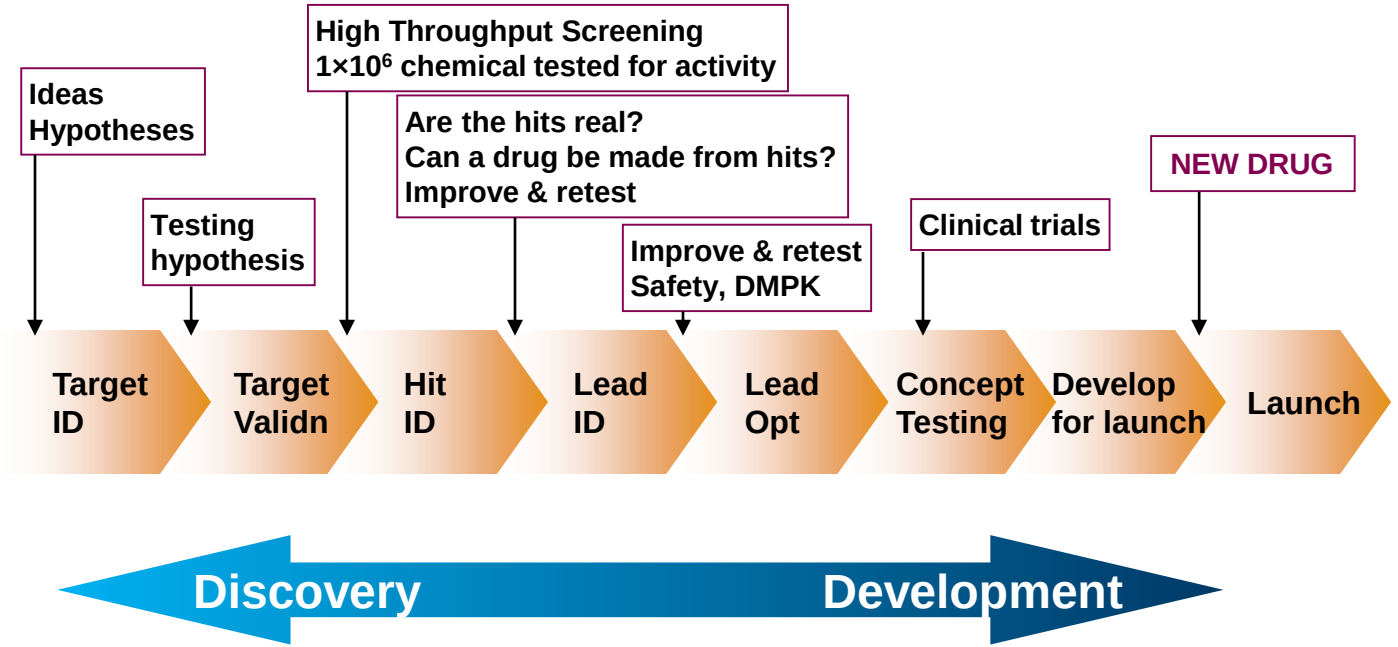
Chronic Cough

18M patients in US

No effective treatments, cough up to 1000x/day

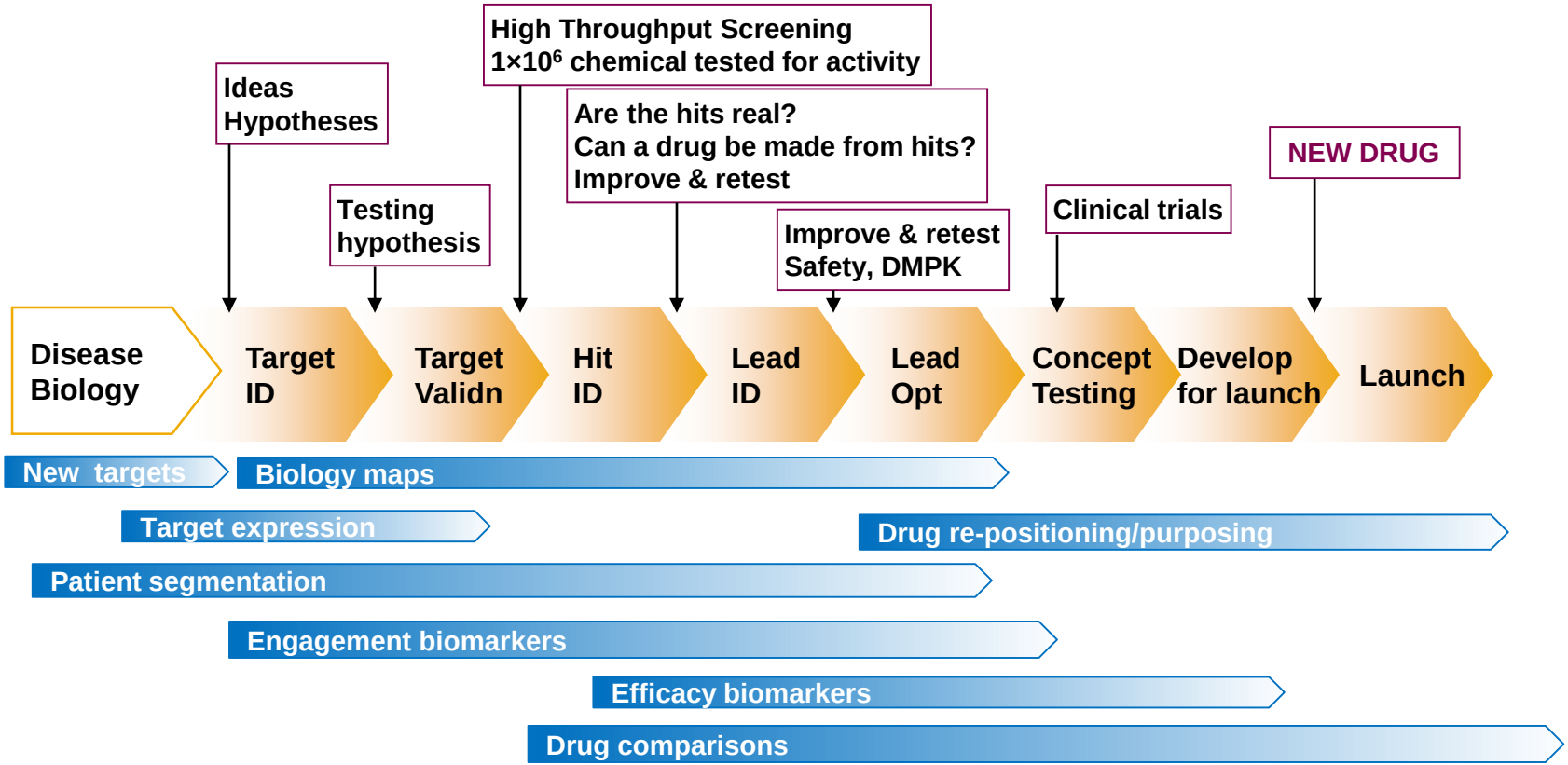
5

The drug discovery process



10-12 years, $2billion+

The drug discovery process

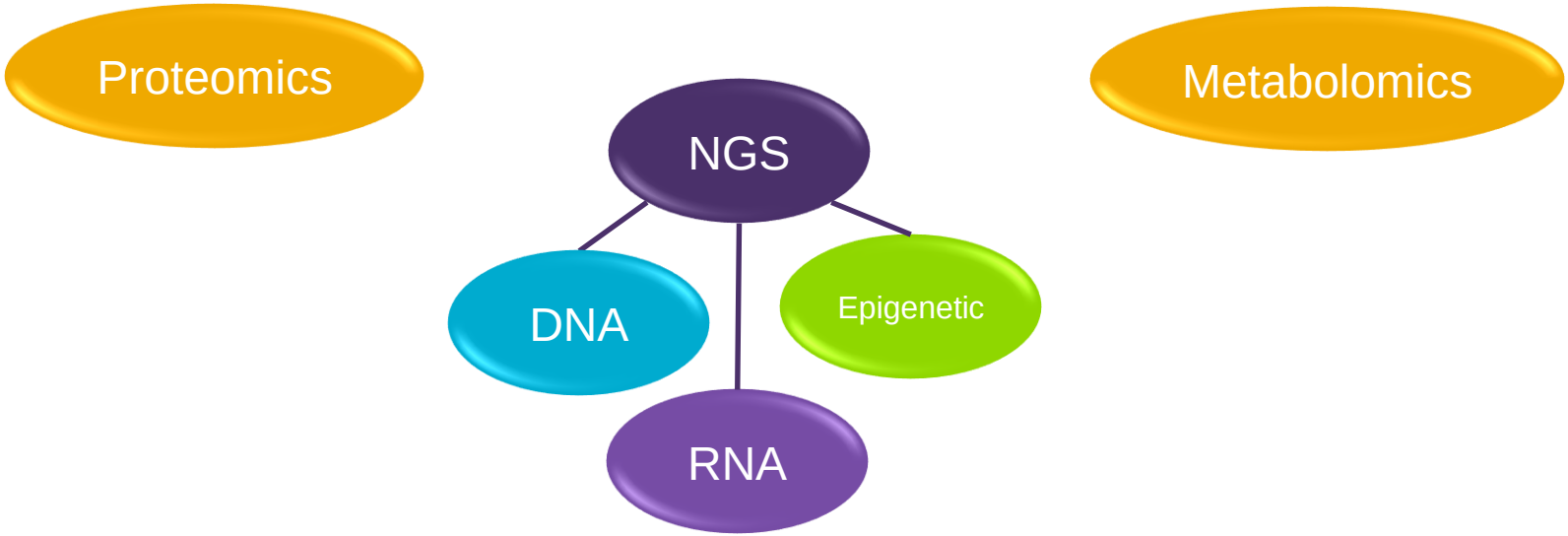


It all starts with a question

- How does our compound behave differently from competitors?
- What are the molecular mechanisms that define disease subtypes?
- What patient group will my drug be most suited to treat?
- Can we identify various cell subtypes in the lung?
- Can we identify microRNAs as biomarkers for IPF?

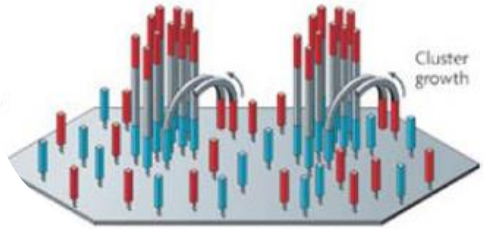
Omics and the importance of hypothesis and question

The hypothesis and question will dictate if an omics technology is the right methodology to use, and if so which one



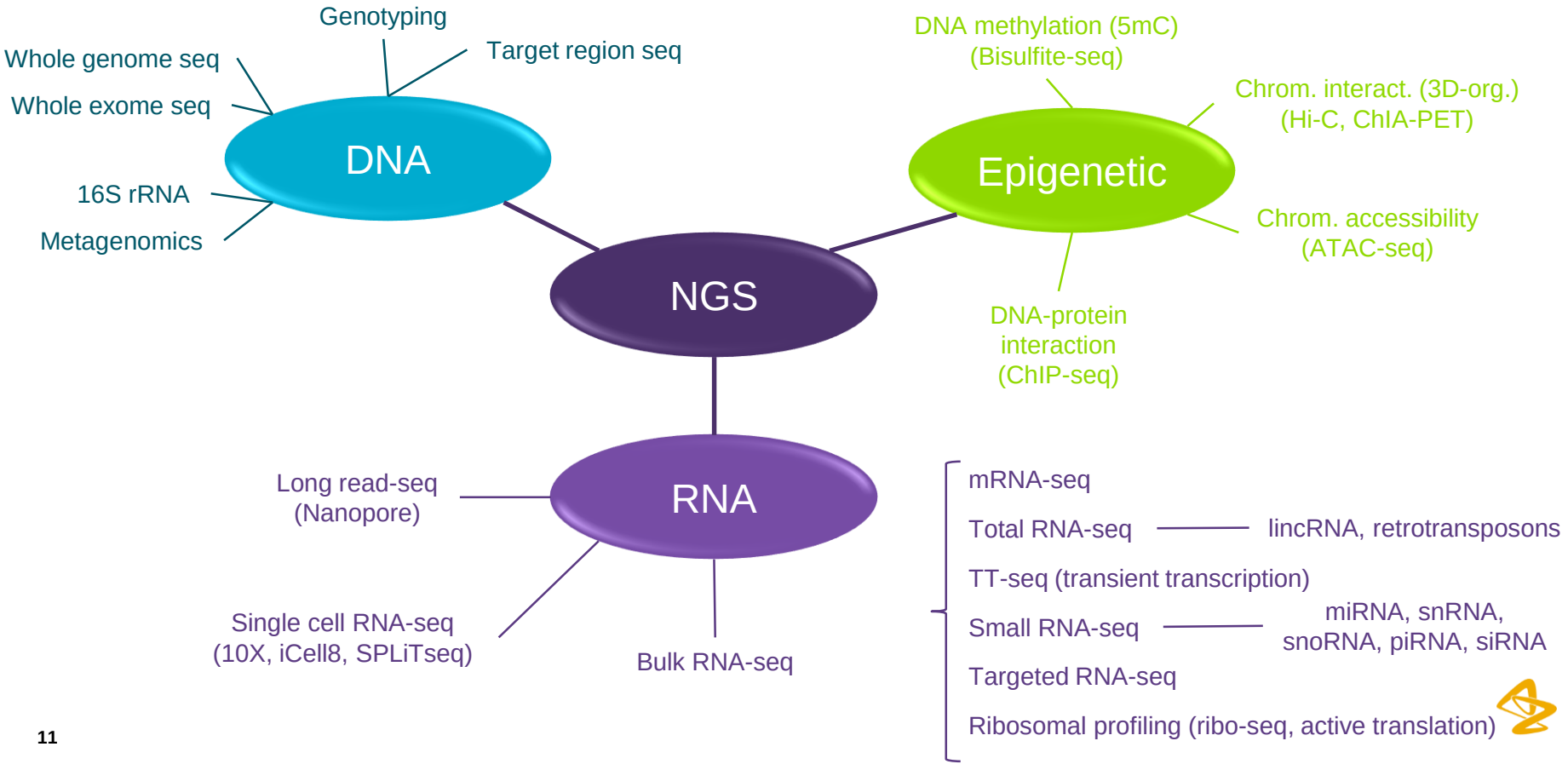
Next Generation Sequencing – a flood of data

- Methods based on the revolutionizing technology of massive parallel sequencing:
  1. Isolate the nucleotide sequences of interest (RNA or DNA)
  2. Prepare library (cDNA, barcode indexes, sequence adapters)
  3. Sequence (Illumina, Ion Torrent, Roche, SOLiD, **Nanopore**)
  4. Map reads to genome (or do de-novo assembly)
  5. Data analysis and interpretation

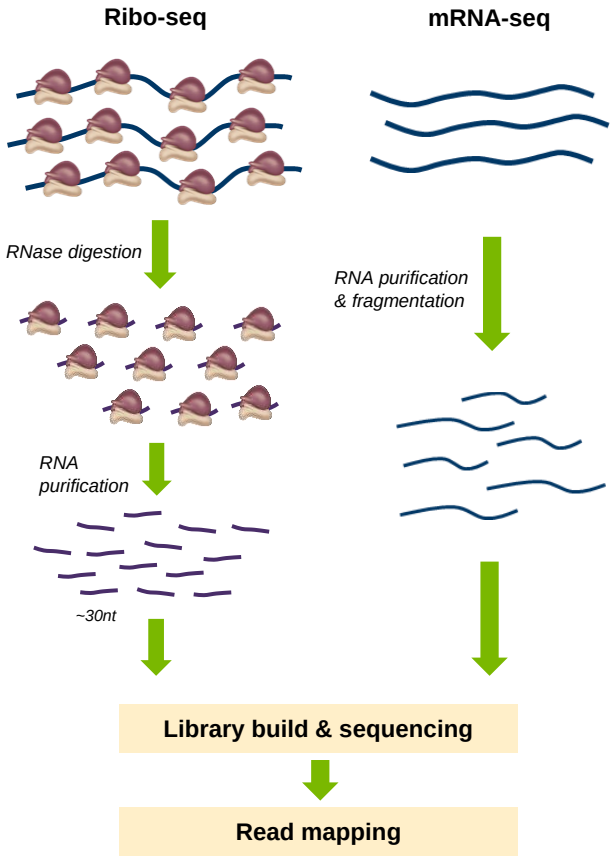


100-200 million molecular clusters

NGS - fast and deep sequencing of DNA or RNA



Ribosome sequencing



- Comprises deep sequencing of **ribosome footprints** (ribosome-protected mRNA fragments).[1]
- Detects **changes in translation** and **better correlates** with downstream protein expression, compared to RNA-seq.[2]
- Provides a snap-shot of global protein translation at **sub-codon resolution.**

Application

The high-resolution, functional signature makes Ribo-seq a potentially valuable tool for:

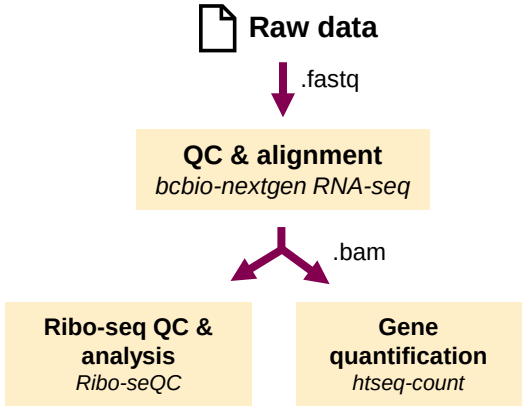
- Strengthening our understanding of disease
- Biomarker discovery & patient stratification

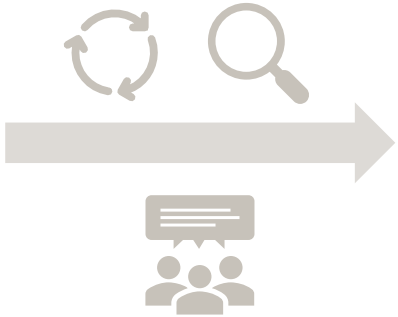
1. Brar *et al*, Nat Rev Mol Cell Biol. 2015 Nov;16(11):651-64
2. Eastman et al, Comput Struct Biotechnol J. 2018 May; 16:167-176.

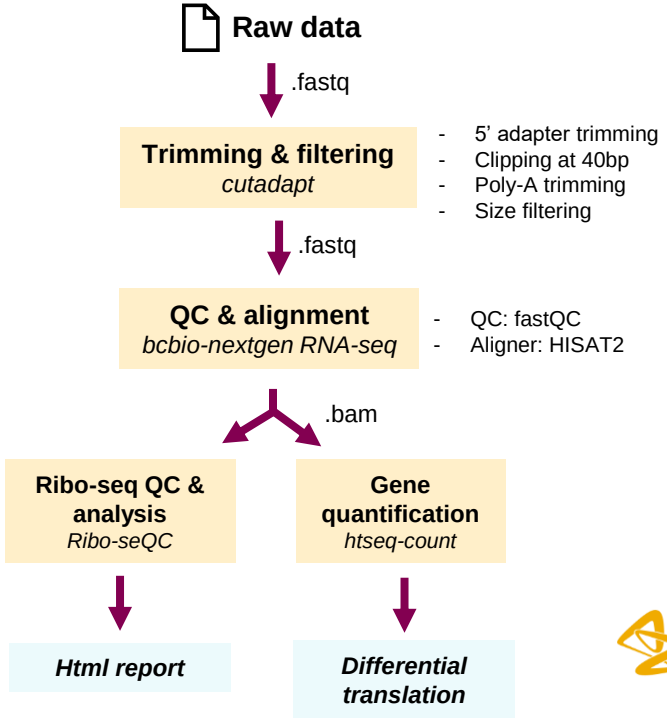
Goal 1. Workflow for Ribo-seq analysis

June | July | August | September | October | November | December | January

R, NGS, Ribo-seq tools → QC optimisation



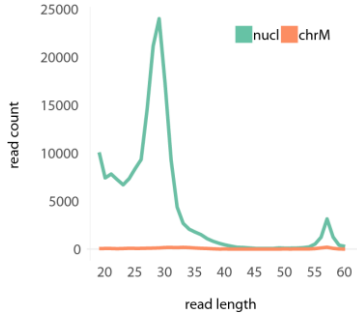




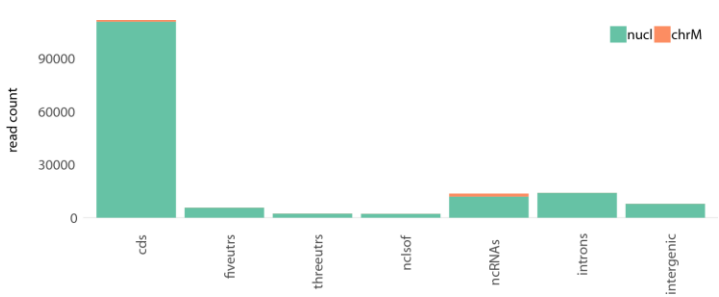
Goal 1. Workflow for Ribo-seq analysis

Ribo-seQC: Ribosome footprint analysis

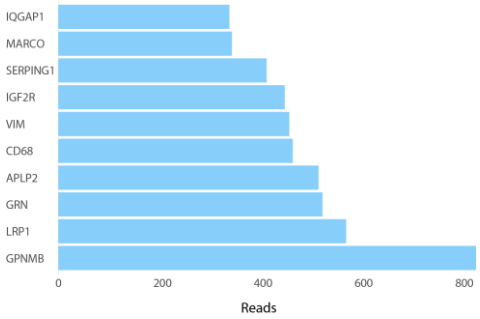
A) Read size distribution



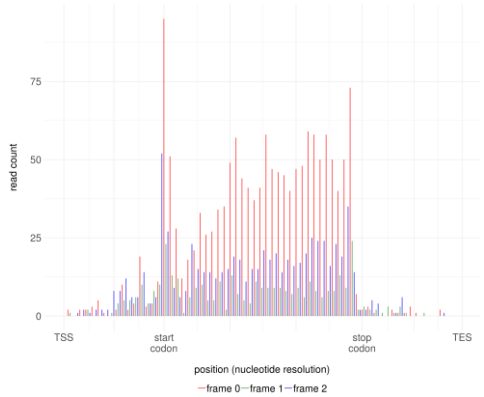
B) Read location distribution



C) Top abundant protein coding genes



D) Ribosome P-site metagene plot

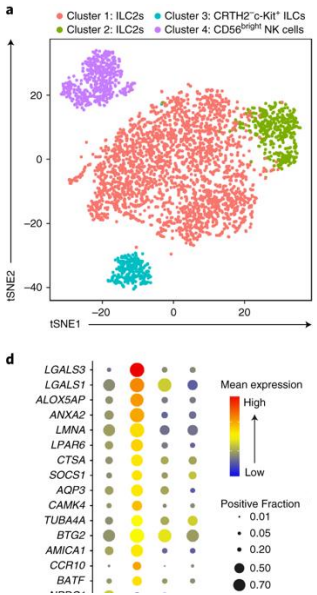


Reads primarily map to **protein coding regions** and the **canonical frame**, are of **expected size** and are enriched in **hallmark macrophage genes.**

Keeping up with technological advances

ILC2 single cell RNA-seq:
Single cell sequencing was applied to characterize the heterogeneity of type 2 innate lymphoid cells (ILC2). A subset of these cells was identified with potential to convert into ILC3-like cells and contribute to the pathology of psoriasis.

This work was published in ***Nature Immunology*** (2019)

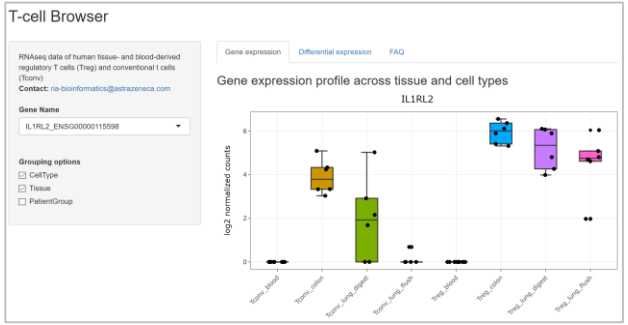


Yoichiro Ohne, Jingya Wang, et al.

https://www.nature.com/articles/s41590-019-0423-0

T-cell (Treg, Tconv) biology and data browser:
Total RNA sequencing was applied to explore the coding and long-non coding transcriptome from human tissue resident and circulating T-cells. Several tissue and cell type-specific genes were identified.

This work was published in **Oncotarget** (2018) and data is available via the **T-cell browser**:



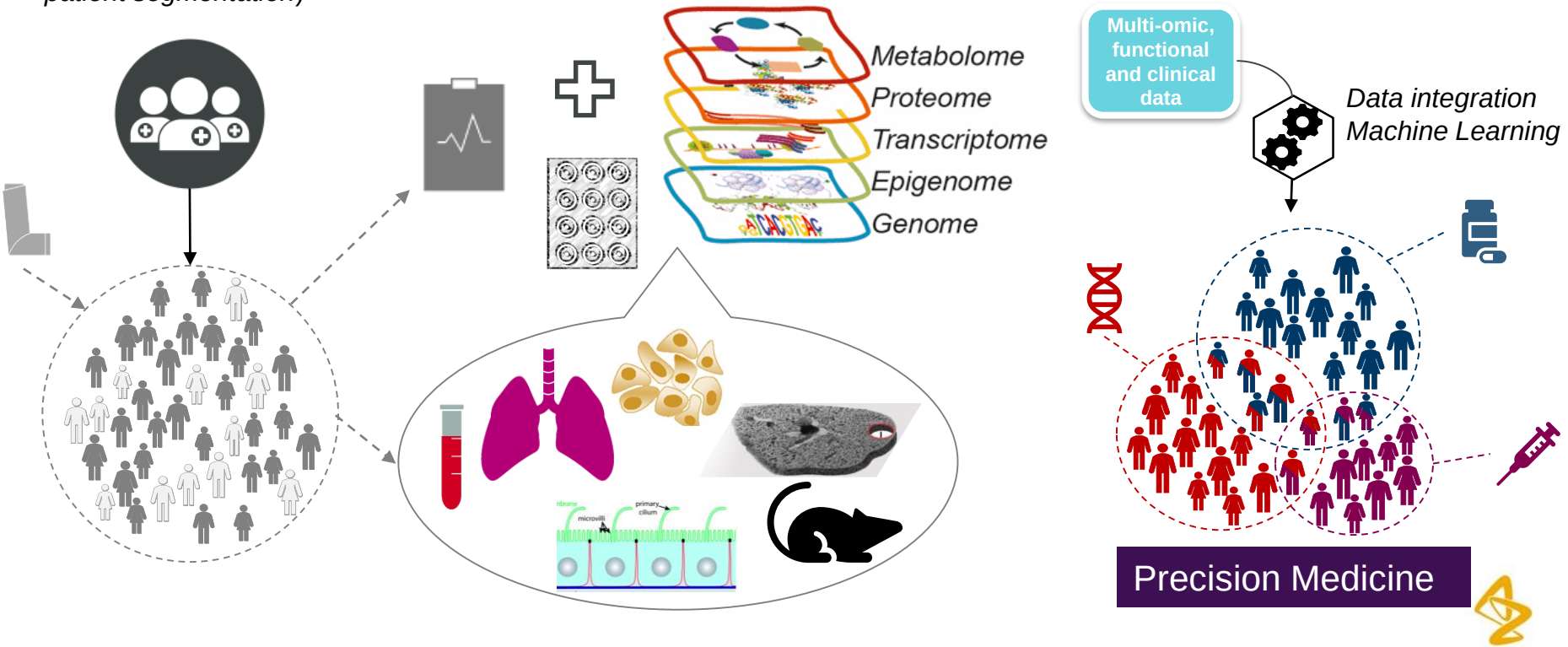
Magda Niedzielska, Lisa Israelsson, Bastian Angermann, et al.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6281418/
https://ria-bioinformatics.shinyapps.io/t_cell_browser/

Tackling Complex Respiratory Diseases

Finding novel targets and Personalising treatment

- COPD, Asthma and IPF are all complex diseases with multiple subtypes (many diseases in one).
- We really want to treat the underlying *individual* molecular defect as opposed to the general symptoms (*precision medicine, patient segmentation*)



AI, Big Data and Digital Health
-bringing the right molecule to the right patient

AstraZeneca recognises that AI/Machine Learning, Big Data and Digital Health have the power to revolutionize healthcare and drug discovery.

- Artificial intelligence (AI) is influencing all stages of the drug discovery and development process. We are taking an AI approach to image analysis, screening assays, drug safety and chemistry (in silico design). AZ recently announced our collaboration with **BenevolentAI** to use machine learning and AI to discover potential new drugs for chronic kidney disease (CKD) and idiopathic pulmonary fibrosis (IPF).

https://azusgb01-cms.visualforce.com/apex/Main?sname=Intranet&name=newsDetailPage&content_id=a0m2A00000ODB1WQAX#/

- AstraZeneca's genomics initiative will analyse genomes from 2 million patients which along with extensive clinical data, will help to understand the influences of genomics on both the causes of disease and responses to treatments.

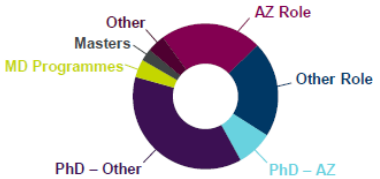
- Efforts are underway to make AZ data more accessible across the organisation. *Need to ensure we do this while still protecting patients rights and privacy.*

Early career opportunities with AstraZeneca

Project work (Masters) - 3 to date in RIA bioinformatics. Many more across the organisation in other functions (mostly wet work).

PhD – case by case

Graduate programme (most relevant ones)



- IMED (all sites) - incorporates three challenging eight month placements across a broad range of IMED functions, supporting drug projects or work relating to IMED's strategic priorities (next intake March/April 2020)
- Data Sciences & AI (all sites) - two-year programme begins with an introduction to drug discovery, then select two eight-month placements where you'll apply your developing skills and experience pushing the frontiers of innovation.

IMED Post-doc programme: Proposals from researchers within AstraZeneca for Post-doc to work on a specific research project/scientific hypothesis they have identified. 48 were funded organisation wide for 2017

Want to know more?

Graduate opportunities at AstraZeneca:
https://careers.astrazeneca.com/students

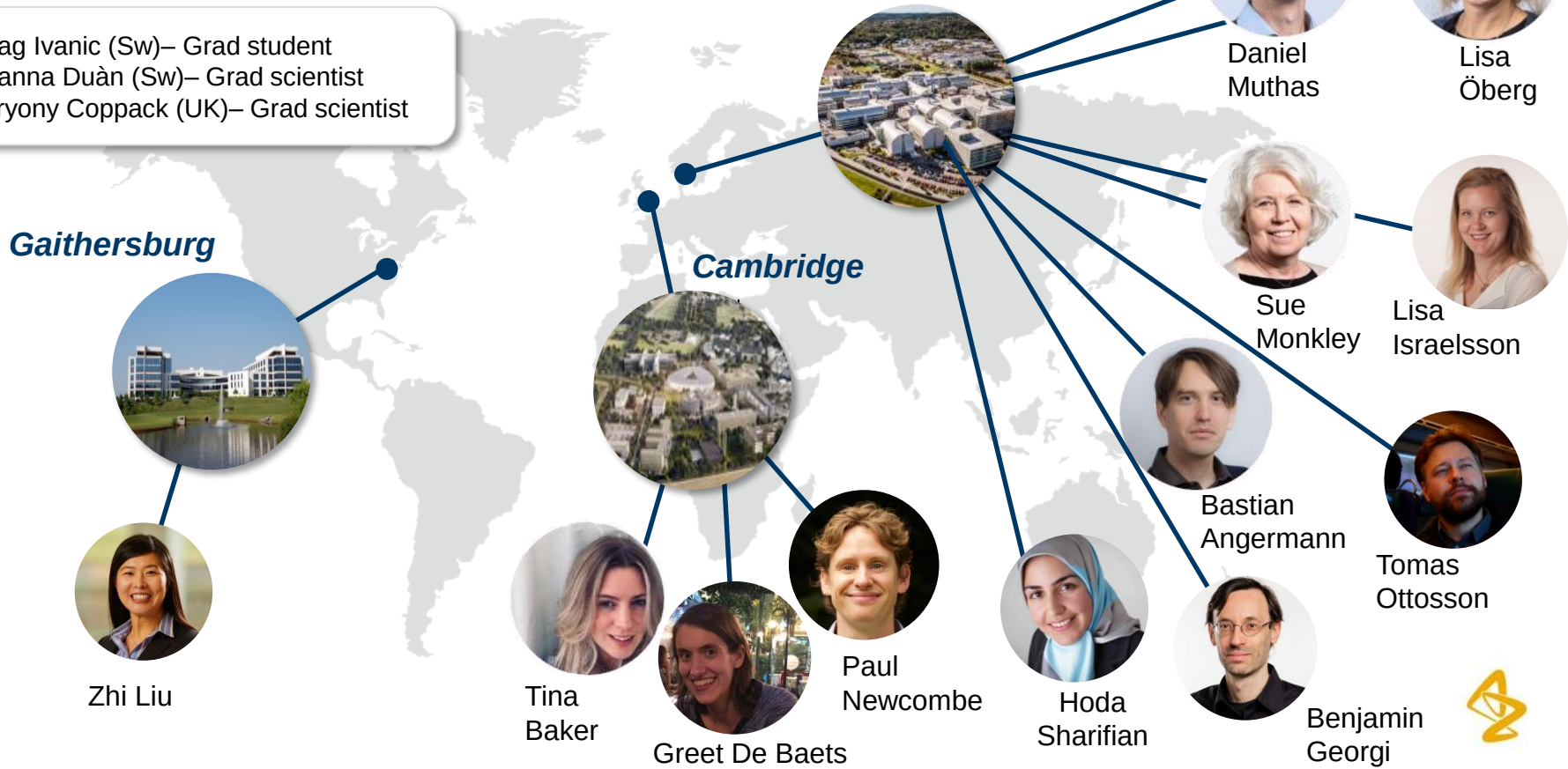
We can arrange a guided tour around AZ Gothenburg that shows what we do and how we do it.

Speakers contacts:

Sue: susan.monkley1@astrazeneca.com

Early R&I Data Science & Bioinformatics

Dag Ivanic (Sw)– Grad student
Hanna Duàn (Sw)– Grad scientist
Bryony Coppack (UK)– Grad scientist



Confidentiality Notice