Genomics england

The 100,000 Genomes Project

Jim Davies

Seattle, October 2015

The 100,000 Genomes Project

Project announced by the Prime Minister in December 2012

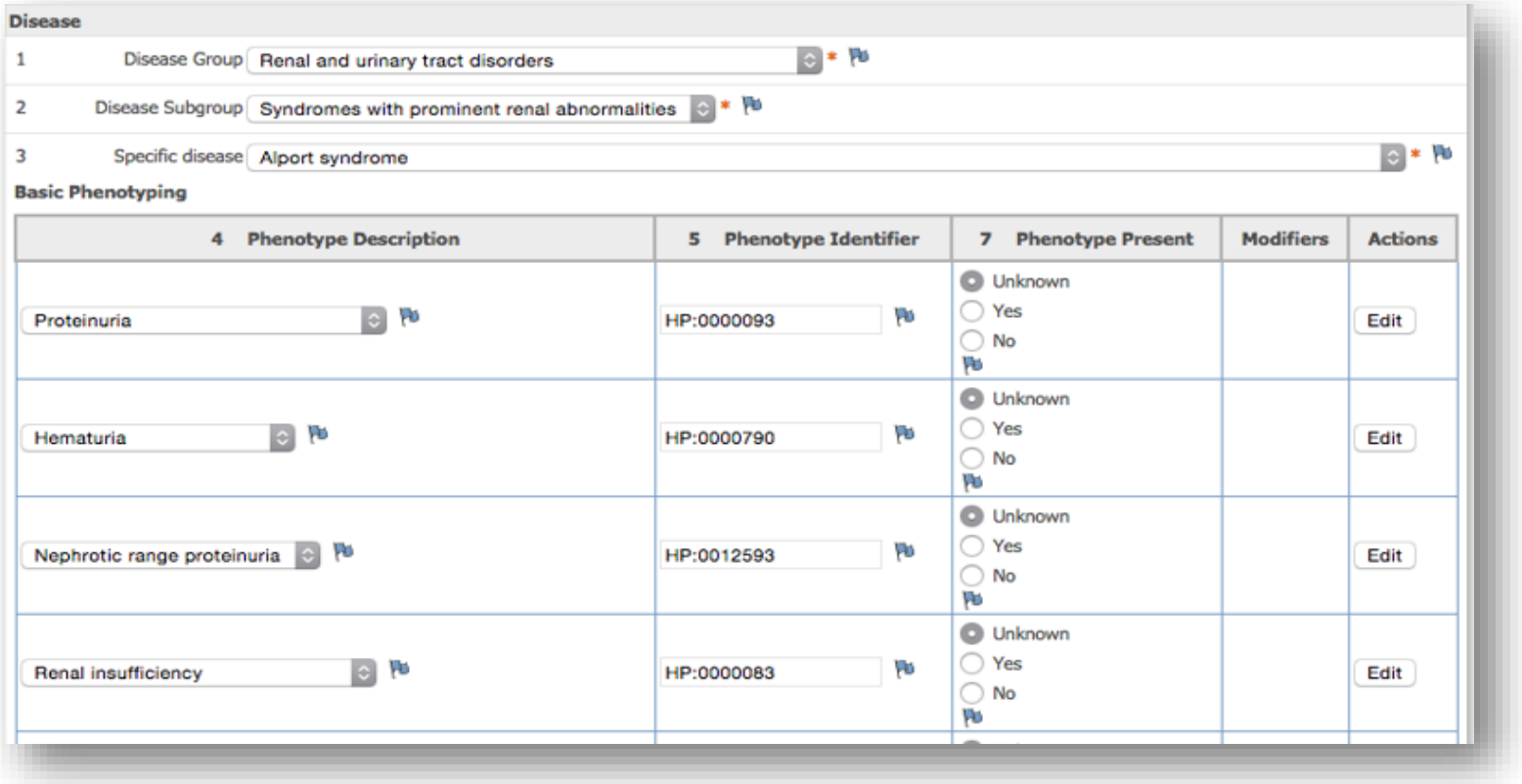
Genomics England announced by Secretary of State for Health in July 2013

Rare Disease

- incidence for each disease < 0.05%
- > 7000 diseases
- cumulative incidence in population ~5%
- cardiovascular, endocrine and metabolism, gastroenterology and hepatology, hearing and sight, immunology and haematology, inherited cancer predisposition, musculoskeletal, neurological, paediatric sepsis, paediatrics, renal, respiratory, skin
- 50,000 participants

Rare Disease

- 30x coverage whole genome sequences
- phenotyping using Human Phenotype Ontology



Disease

1 Disease Group: Renal and urinary tract disorders *

2 Disease Subgroup: Syndromes with prominent renal abnormalities *

3 Specific disease: Alport syndrome *

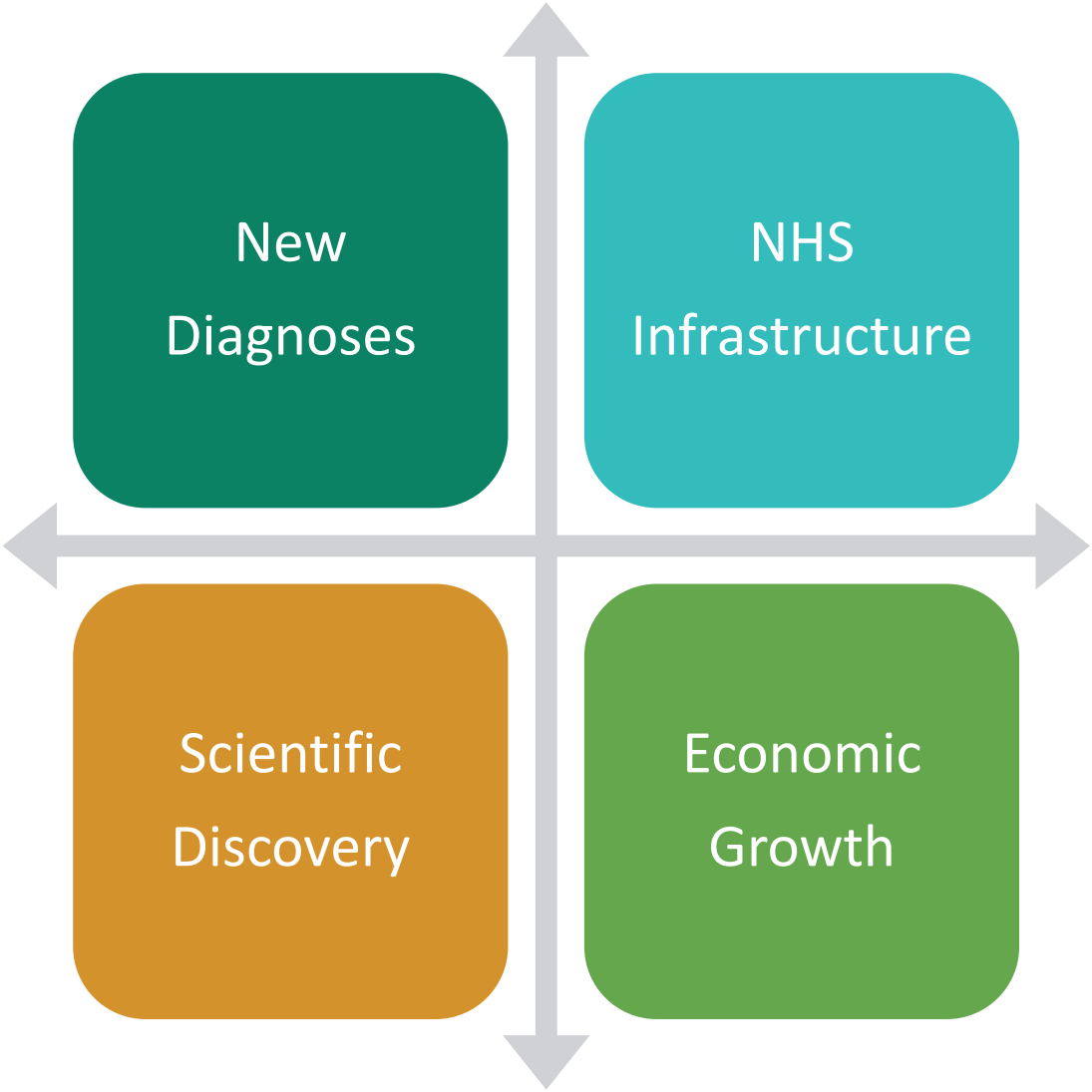
Basic Phenotyping

4 Phenotype Description	5 Phenotype Identifier	7 Phenotype Present	Modifiers	Actions
Proteinuria	HP:0000093	Unknown / Yes / No		Edit
Hematuria	HP:0000790	Unknown / Yes / No		Edit
Nephrotic range proteinuria	HP:0012593	Unknown / Yes / No		Edit
Renal insufficiency	HP:0000083	Unknown / Yes / No		Edit

Cancer

- ovarian, breast, colorectal, prostate, lung, leukaemia, sarcoma, renal, rare and childhood cancers, cancer of unknown primary
- 25,000 participants
- 75x coverage somatic, 30x coverage germline
- dataset reflects requirements for genomic medicine

Expectations



Participants

NHS England

NHS Genomic Medicine Centres

Consented participants, samples, phenotyping, clinical and laboratory data, clinical reports, confirmatory testing, and follow-up

Biorepository

biocentre uk
The centre for biomedical services

DNA & samples for multi-omics

Sequencing

illumina®

Data

Genomics england

UNIVERSITY OF OXFORD

Identified Data

Collected against agreed models, linked to secondary data from other sources.

Other Data

Hospital Episode Statistics, registry data, mortality, *and more*

Genomics england

Farr

Research Data

De-identified versions of the data made available via a virtual desktop infrastructure

Tools and Services

Quality assurance, clinical interpretation, tools for analysis, reference data

Researchers

Researchers

Genomics England Clinical Interpretation Partnerships, Industry Consortia, and other approved researchers

Infrastructure

- Genomic Medicine Centres (NHS England)
- NIHR national biosample centre
- Hinxton sequencing centre (Wellcome Trust)
- secure data centre (NIHR, MRC)

biosample centre

sequencing centre

data centre

Data Acquisition

- registration, consent, family history, phenotyping via web-based interface
- clinical and laboratory data submitted as XML messages, reports upon clinical events
- secondary data from clinical audits and national reporting datasets
- integrated, longitudinal record via LabKey server for review and quality assurance

- patient-reported data

Research Environment

- separate data centre
- a research ‘embassy’ for each group
- access to shared LabKey server
  - explore data within policy
  - take snapshots for analysis
- private cloud
- ‘airlock’ system

- virtual desktop infrastructure

Progress

- 11 Genomic Medicine Centres
- 5000 whole genome sequences
- 125 rare diseases
- 5 cancers
- 1 pathogen
- 4 tool providers
- 10 pharma companies

- 2500 researchers

Contributors (include)

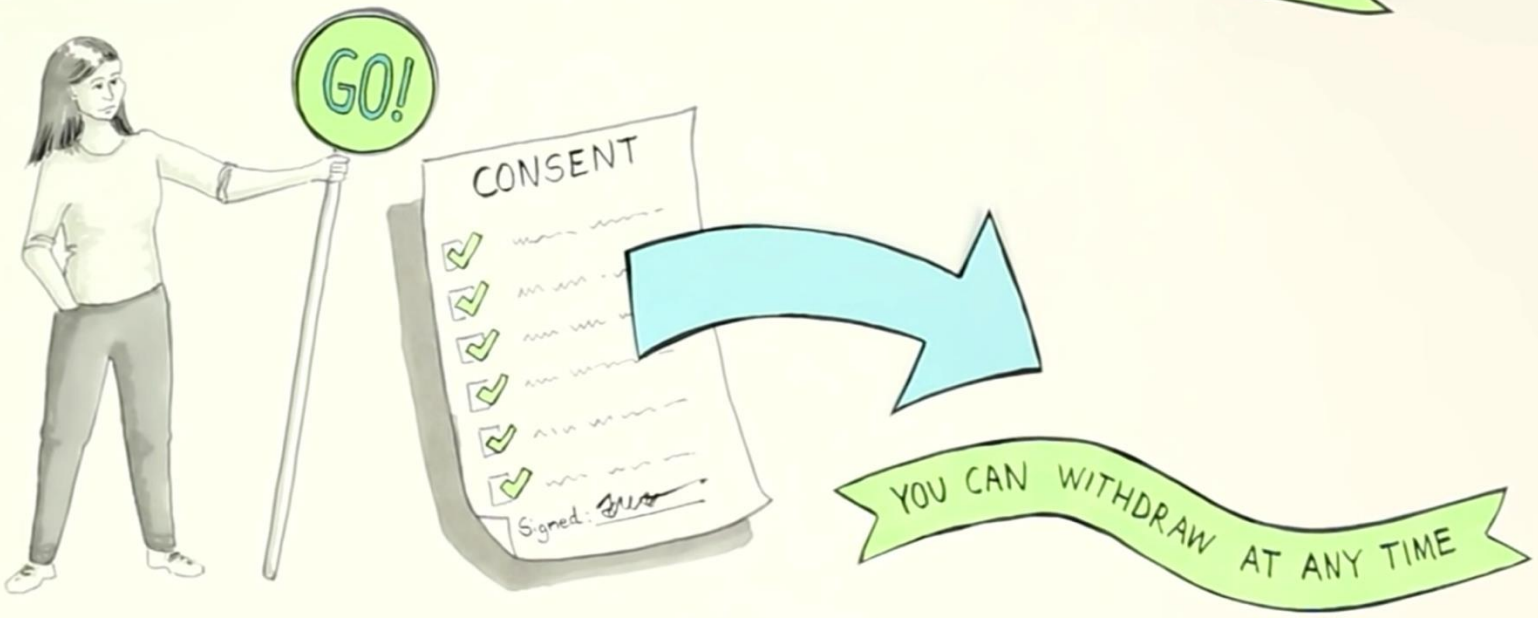
- NIHR Biomedical Research Centres
- NIHR Bioresource for Rare Diseases
- NIHR Health Informatics Collaborative
- NHS England, Public Health England
- Medical Research Council
- Cancer Research UK
- Wellcome Trust
- ...

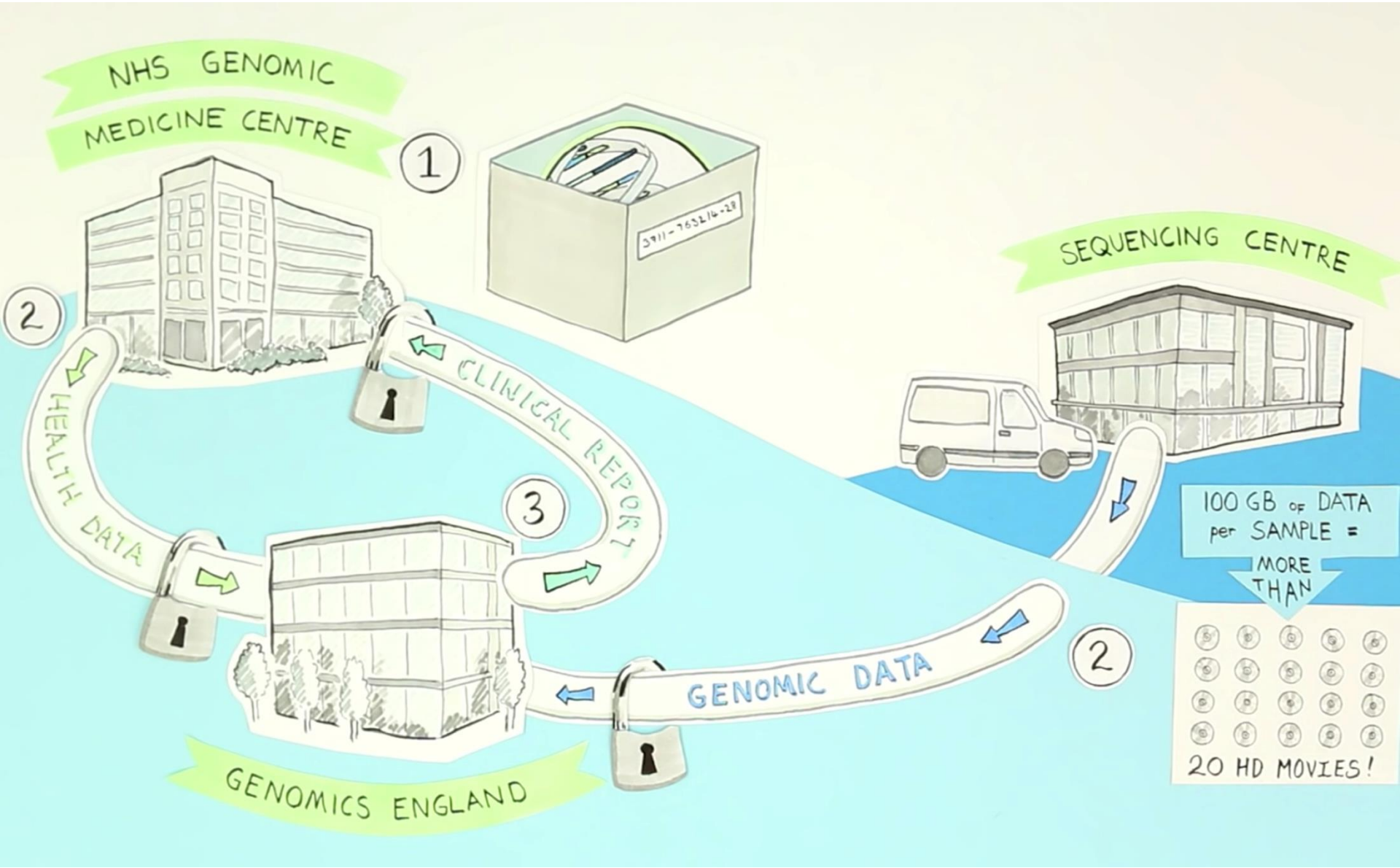
Challenges

- earn the trust of:
  - patients
  - clinicians
  - researchers

- help to transform:
  - clinical informatics
  - research informatics
  - culture

WHAT DOES GENOMICS ENGLAND DO WITH YOUR DATA ?




NHS GENOMIC MEDICINE CENTRE
1
2
HEALTH DATA
CLINICAL REPORT
3
GENOMICS ENGLAND
GENOMIC DATA
2
SEQUENCING CENTRE
100 GB OF DATA per SAMPLE =
MORE THAN
20 HD MOVIES!

3

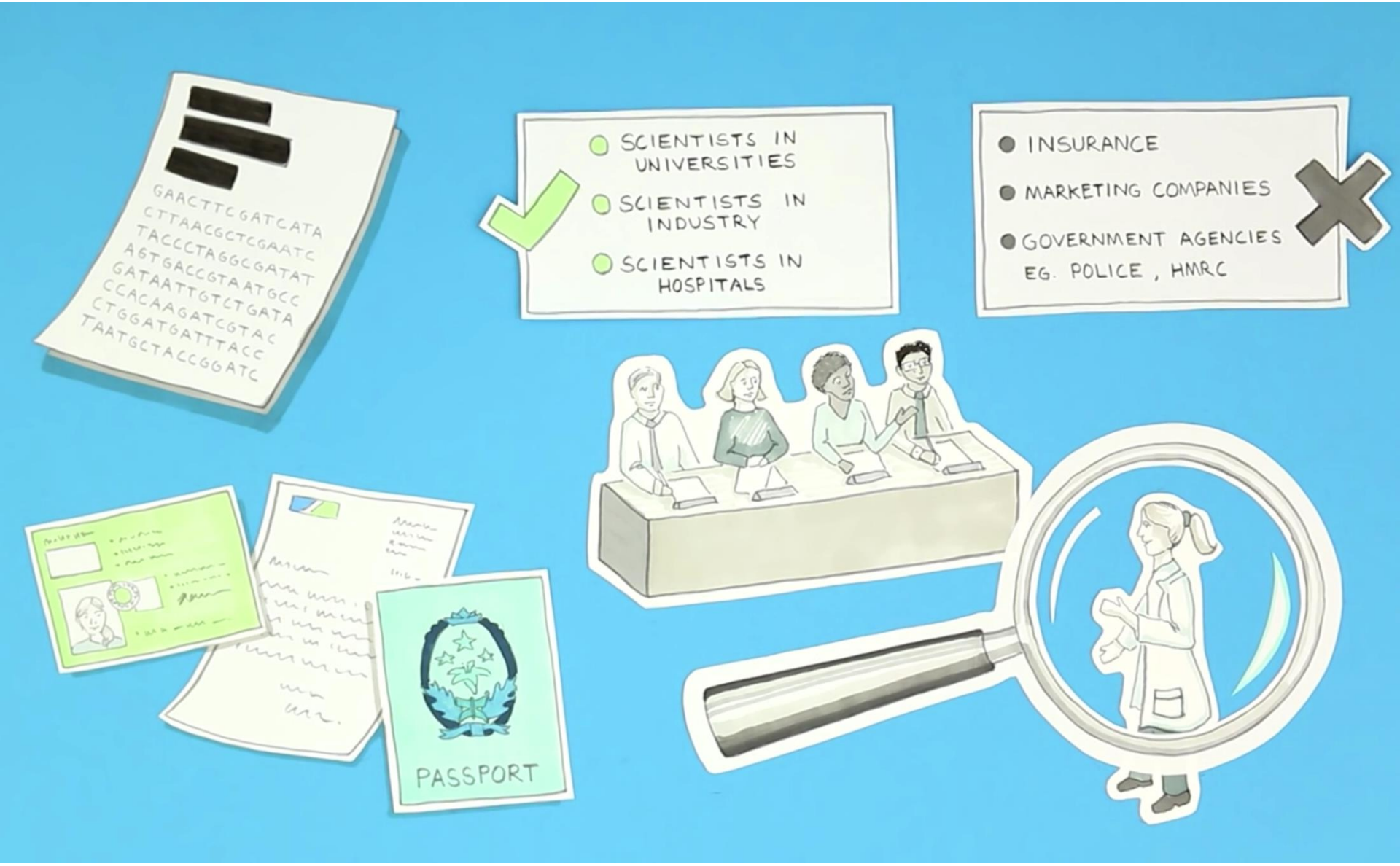
- YOUR HEALTH DATA
- YOUR MEDICAL RECORD DATA
- YOUR GENOME DATA

WHAT WILL SCIENTISTS DO WITH THE DATA ?

UNDERSTAND WHICH DIFFERENCES ARE IMPORTANT

BEST TREATMENT

- USE CURRENT MEDICINES BETTER
- NEW DRUGS
- NEW DIAGNOSTIC TESTS


GAACTTCGATCATA
CTTAACGCTCGAATC
TACCCTAGGCGATAT
AGTGACCGTAATGCC
GATAATTGTCTGATA
CCACAAGATCGTAC
CTGGATGATTTACC
TAATGCTACCGGATC
SCIENTISTS IN UNIVERSITIES
SCIENTISTS IN INDUSTRY
SCIENTISTS IN HOSPITALS
INSURANCE
MARKETING COMPANIES
GOVERNMENT AGENCIES EG. POLICE , HMRC
PASSPORT

100,000 GENOMES PROJECT

Genomics england

YOUR DATA MAKES A DIFFERENCE

Thanks

- The Genomics England team: in particular,
  - Sir John Chisholm (Executive Chair)
  - Professor Mark Caulfield (Chief Scientist)
  - Vivienne Parry (Communications and Outreach)
  - Ed Stafford (Director of Informatics)

- LabKey
  - enjoy the rest of the conference!