Research for the Consumer, Powered by the Consumer

Mike Polcari, Joyce Tung

Who is 23andMe?

- **Direct to consumer genetic testing company**
- Founded in 2006 by Linda Avey and Anne Wojcicki
- Customers receive uninterpreted raw genotype data and genetic results on their ancestry
- Granted authorization by FDA to market a direct-to-consumer genetic test (Bloom syndrome)
- 1M+ genotyped customers

23andMe

Our Mission:

To help people access, understand and benefit from the human genome.

How can people benefit?

What can we gain from research into human genetics?

- Help find the **right treatments for the right people**
  - Use genetics to personalize treatment
  - Understand risk to maximize **prevention**

- Help develop **better drugs** using genomic data from **everyday people**

23andMe

The Research Platform

How do people get started?

1 Order a kit online

2 Provide your DNA Sample

3 Register your sample

4 Mail your kit back

23andMe

Consenting to participate

- Research protocol and consent document are approved by an **external ethics review board** (IRB)
- People make an **informed decision** to contribute their data to research
- Their genotype and phenotype data will be used in a **broad range of studies**
- **80%** of customers consent to participate in research

23andMe

Advantages of online research

- **Participation is easy**
  - People can get involved at any time of day, from their couch

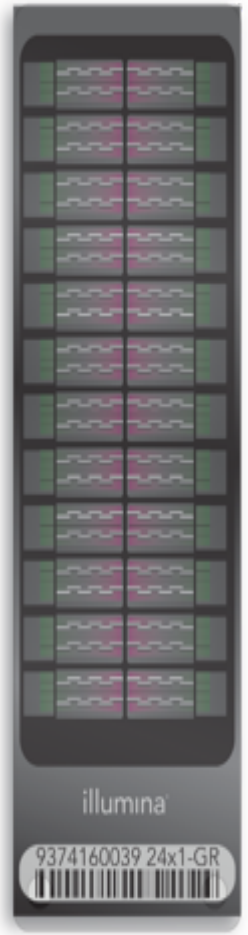
- **Geography is not a barrier**
  - Proximity to a clinical study site is not a requirement

- **Everyone can be in multiple studies at once**
  - The average 23andMe participant is genotyped only once, but is part of 200+ studies

23andMe

Genotype Data

- Currently based on a **fully customized Illumina HumanOmniExpress+ Beadchip**, with SNPs selected to provide good genome-wide coverage for imputation and coverage of variants of known or suspected medical relevance
- Currently imputing with 1000G phase 1, looking into 1000G phase 3 and UK10K
  - Phase with Beagle, impute with Minimac
- IBD calculated by HAPLOSCORE
- Ancestry estimated using our Ancestry Composition algorithm



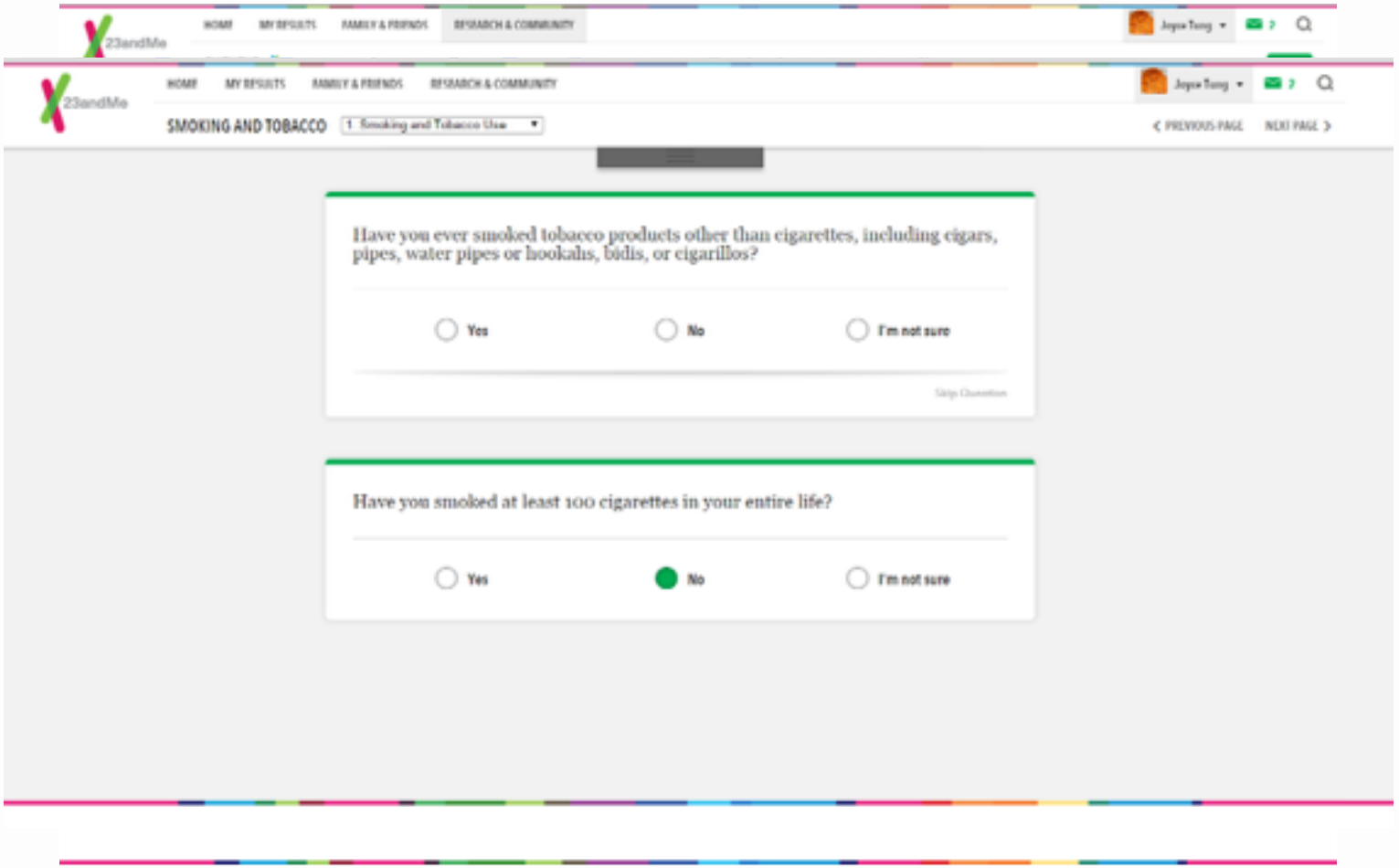
23andMe

Phenotype Data

- Specially trained survey methodologists develop **online surveys** to collect high quality phenotypic data for research
- Topics range from **eye color** to **caffeine consumption** to **Parkinson's disease**
- Customers can **choose** to take as many surveys as they want whenever they want
- Now **thousands** of questions asked, organized into over a thousand "phenotypes" for analysis
- Over **250 million data points** collected

23andMe

Online Research Surveys

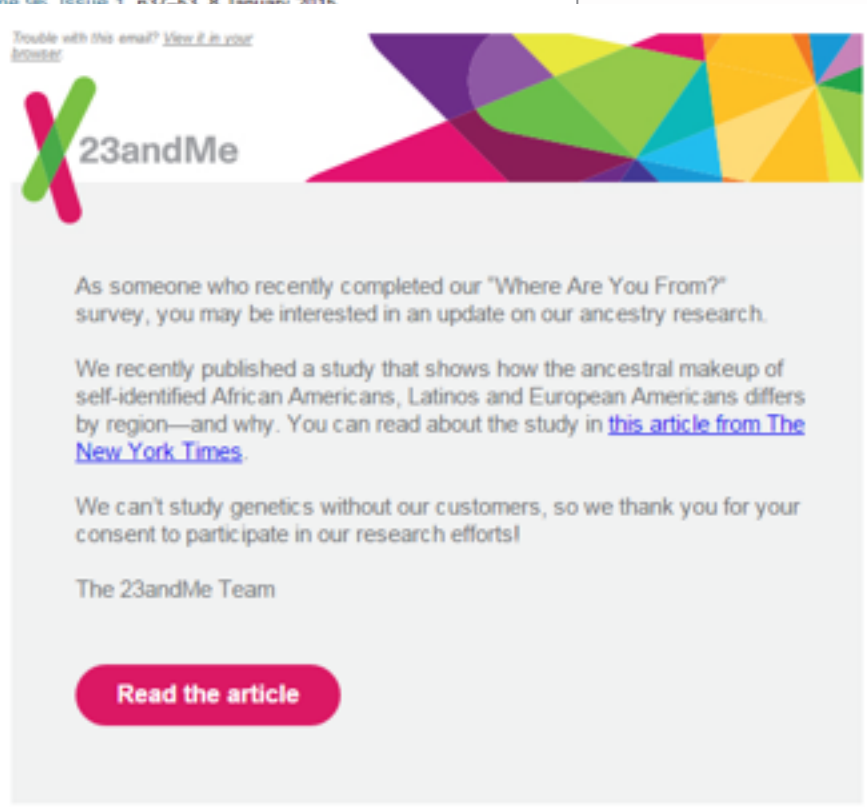


Quality of Self-Reported Data

- Adhere to **best practices** in survey methodology
- **Replicated** genetic associations at hundreds of loci for 50+ conditions
- **Large sample sizes** help mitigate challenges of misclassification, though data quality will always be an important consideration
- Data are structured for research and collected all in the same way

23andMe

A relationship with our participants

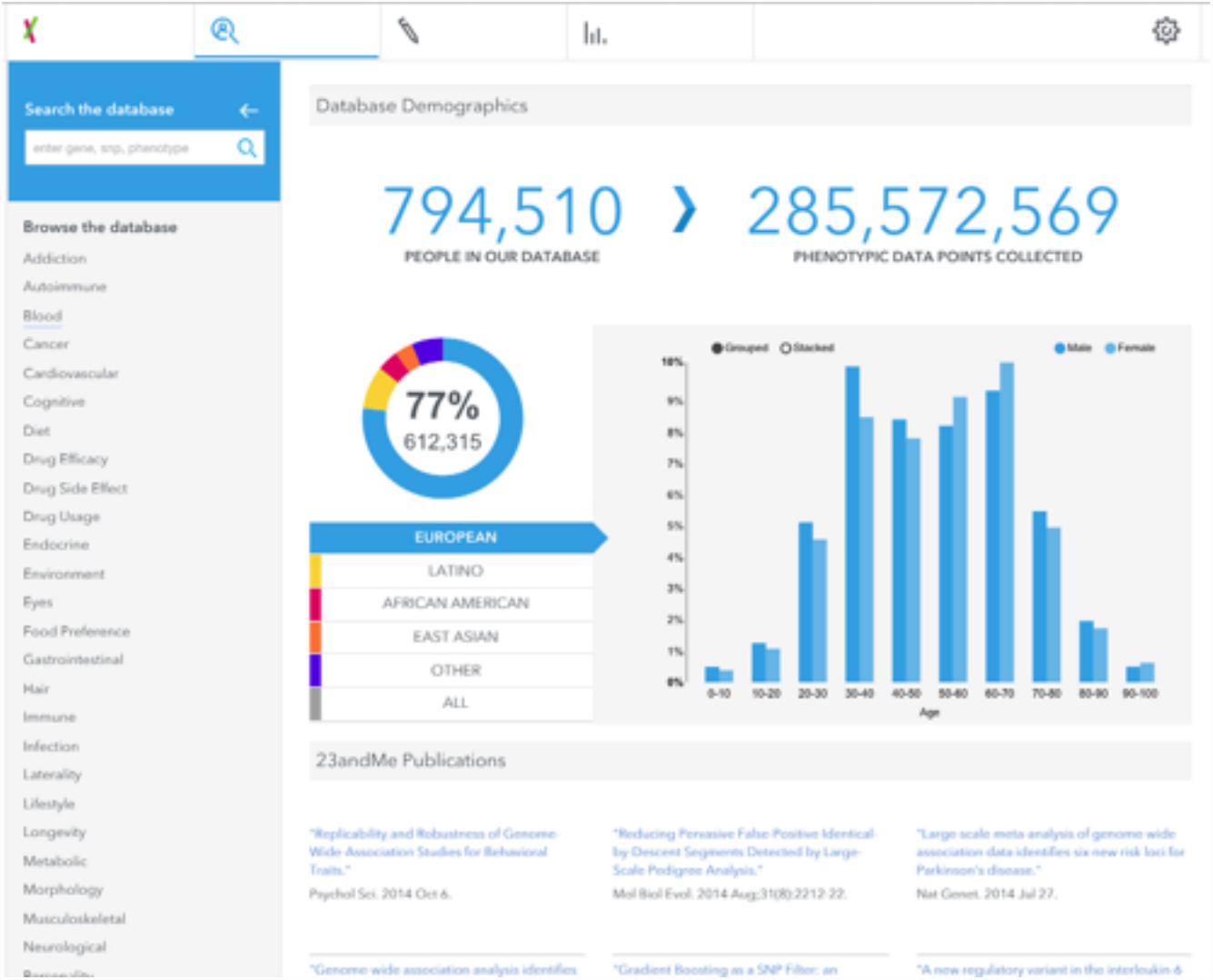


23andMe

API

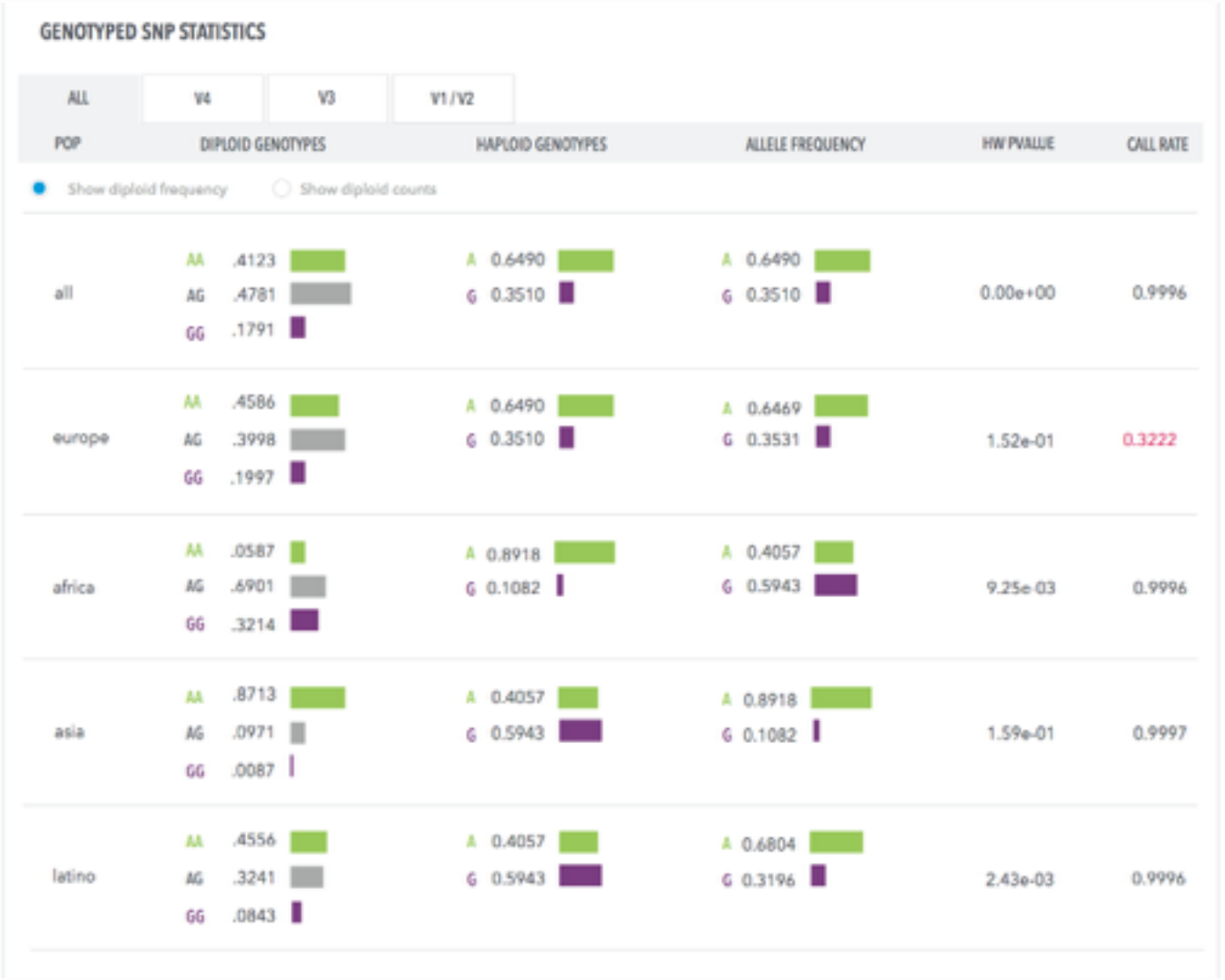
- Customers **may** use OAuth2 to authorize 3rd parties
- 3rd parties include: EMR's, PHR's, Research Studies, Apps…
- Flexible, **secure** REST API includes access most account data: raw genetics, higher-level interpretations, computed ancestry, relationships between users, PII, phenotypes, etc.
- 3rd parties request only relevant subset of data
- api.23andme.com

Research Portal - Overview

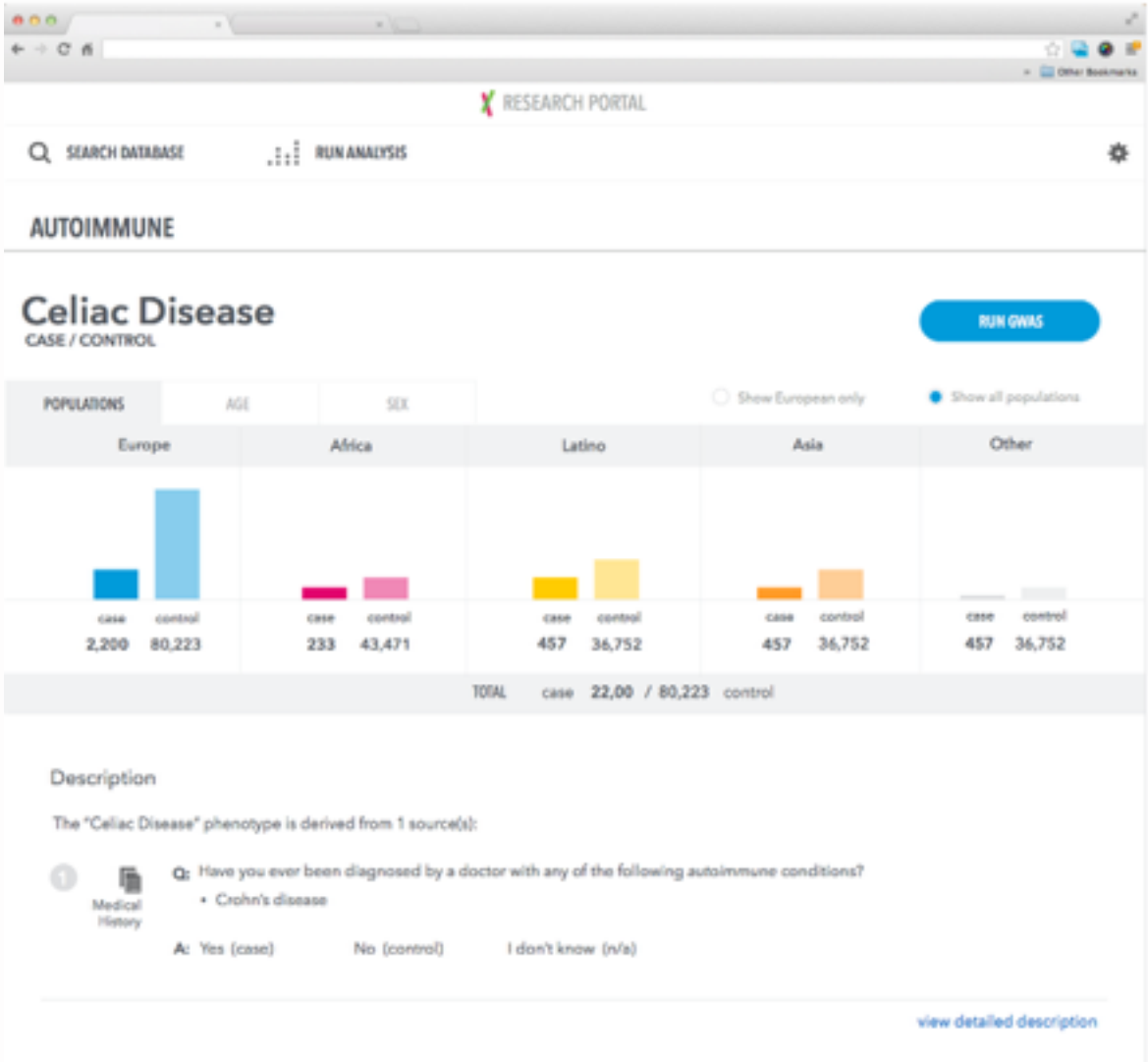


23andMe

Research Portal - SNP



Research Portal - Phenotype



23andMe

Publications and Collaborations

30+

Total publications

40+

Academic collaborations

www.23andme.com/for/scientists

23andMe

NIH Grants

- SBIR Catalyst (NHGRI): Web-based Phenotyping for Genome Wide Association Studies of Drug Response (2010) 1R43HG005807-01
- SBIR Phase I (NHGRI): Development of DNA Sequence Data-Quality Metrics for Personal Genomics (2012) 1R43HG006976-01
- SBIR Phase I (NHLBI): Genetics of Allergic Disease in a Participatory Research Cohort (2012) 1R43HL115873-01
- SBIR Phase I (NHGRI): Development of a web-based database and research engine for genetic discovery (2012) 1R44HG006981-01
  - Phase II: 2013 @R44HG006982-02

23andMe