Introduction
○○○○○○○○○○

Open GWAS
○○○○○

Privacy & Implications
○○○○○○○

Discussion
○○○

Crowdsourcing Genome Wide Association Studies

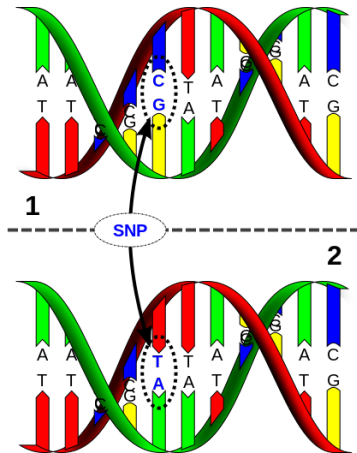Bastian Greshake and Philipp Bayer

28.12.2011

Overview

What are GWAS?

- Genome-wide Association Studies
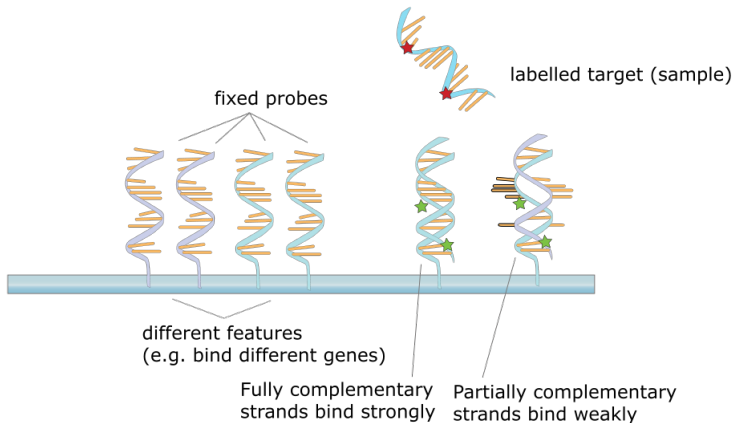
What are GWAS?

- Genome-wide Association Studies
- Link genetic variants (SNPs) to certain traits like eye or hair colour or to diseases like Diabetes, types of cancer

Single Nucleotide Polymorphism

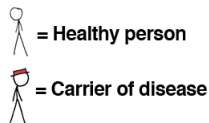Source: http://en.wikipedia.org/wiki/File:Dna-SNP.svg

Introduction
○○●○○○○○○○○

Open GWAS
○○○○○

Privacy & Implications
○○○○○○○

Discussion
○○○

Association studies?

How to analyse SNPs?

Source: http://en.wikipedia.org/wiki/File:NA_hybrid.svg

Introduction	Open GWAS	Privacy & Implications	Discussion
○○○●○○○○○○	○○○○○	○○○○○○○	○○○

Association studies?

How do GWAS work?

= Healthy person

= Carrier of disease

Introduction
○○○○●○○○○○

Open GWAS
○○○○○

Privacy & Implications
○○○○○○○

Discussion
○○○

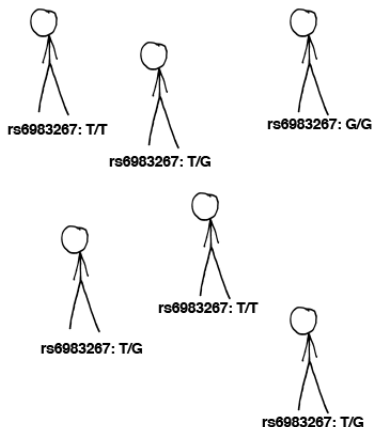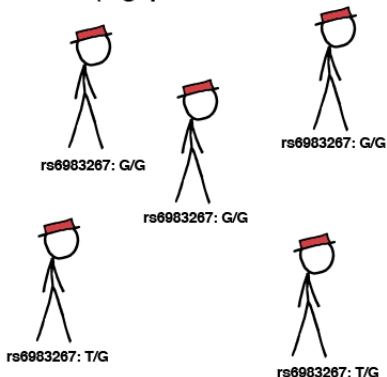Association studies?
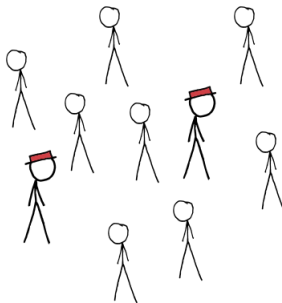
How do GWAS work?

Healthy people

People w/ disease (e.g. prostate cancer)
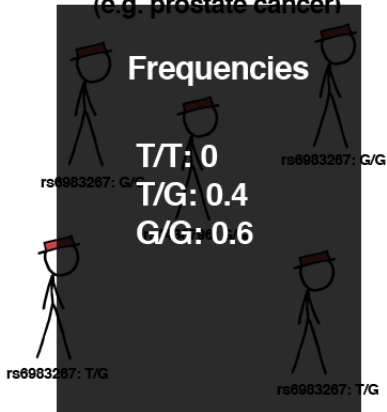
How do GWAS work?
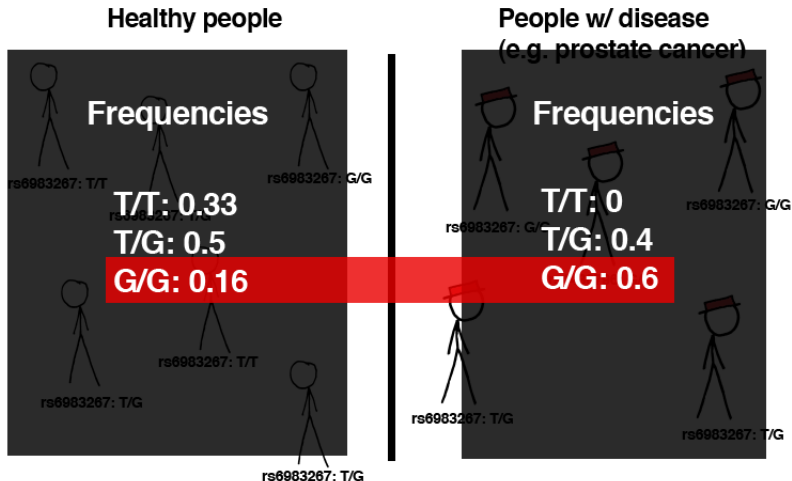
Introduction
0000000●000
Open GWAS
00000
Privacy & Implications
0000000
Discussion
000
Association studies?

How do GWAS work?

Introduction
0000000●00
Open GWAS
00000
Privacy & Implications
0000000
Discussion
000
Association studies?

Some GWAS-examples

- Sladek *et al.* (2007) identified four gene locations linked to heightened type 2 diabetes risk

Introduction
○○○○○○○●○○

Open GWAS
○○○○○

Privacy & Implications
○○○○○○○

Discussion
○○○

Association studies?

Some GWAS-examples

- Sladek *et al.* (2007) identified four gene locations linked to heightened type 2 diabetes risk
- Kogan *et al.* (2011) linked rs53576 (G:G) to pro-social behaviour

Introduction
0000000●00

Open GWAS
00000

Privacy & Implications
0000000

Discussion
000

Association studies?

Some GWAS-examples

- Sladek *et al.* (2007) identified four gene locations linked to heightened type 2 diabetes risk
- Kogan *et al.* (2011) linked rs53576 (G:G) to pro-social behaviour
- The Wellcome Trust Case Control Consortium (2007) linked 24 locations to 7 major diseases

Introduction
○○○○○○○○○●○

Open GWAS
○○○○○

Privacy & Implications
○○○○○○○

Discussion
○○○

Association studies?

Problems with GWAS

- Large enough sample size

Problems with GWAS

- Large enough sample size
- Correcting for multiple testing

Introduction
○○○○○○○○○●○

Open GWAS
○○○○○

Privacy & Implications
○○○○○○○

Discussion
○○○

Association studies?

Problems with GWAS

- Large enough sample size
- Correcting for multiple testing
- Correlation != Causation

Putting GWAS to use

- Direct-To-Consumer genetic testing
- Analyse about 1 million SNPs and provide summary of disease risks & ancestry
- About $200 for a genotyping

Putting GWAS to use

- Direct-To-Consumer genetic testing
- Analyse about 1 million SNPs and provide summary of disease risks & ancestry
- About $200 for a genotyping
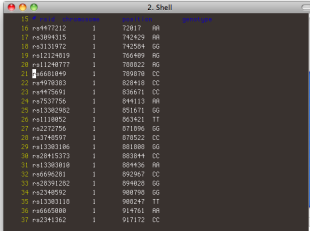- Providers: 23andMe, deCODEme, FamilyTree DNA, ...

Putting GWAS to use

- Direct-To-Consumer genetic testing
- Analyse about 1 million SNPs and provide summary of disease risks & ancestry
- About $200 for a genotyping
- Providers: 23andMe, deCODEme, FamilyTree DNA, ...
- You get access to the raw data!

Numbers on DTC

- 23andMe alone has over 100.000 customers

Numbers on DTC

- 23andMe alone has over 100.000 customers
- 76 % of their customers agree to participate in research

Numbers on DTC

- 23andMe alone has over 100.000 customers
- 76 % of their customers agree to participate in research
- 59 % of them share phenotypic information with 23andMe

Research in company labs

- 23andMe published results of studies with up to 30.000 participants

Research in company labs

- 23andMe published results of studies with up to 30.000 participants
- Replication of older GWAS

Research in company labs

- 23andMe published results of studies with up to 30.000 participants
- Replication of older GWAS
- Finding new associations for Parkinsons disease

Data sharing

- People are already sharing the raw data of DTC tests

Data sharing

- People are already sharing the raw data of DTC tests
- 1-5 % of 23andMe customers would be enough to perform simple GWAS

Data sharing

- People are already sharing the raw data of DTC tests
- 1-5 % of 23andMe customers would be enough to perform simple GWAS
- The *Personal Genome Project*: Open data, but closed participation

Introduction
○○○○○○○○○○○

Open GWAS
○○○●○○

Privacy & Implications
○○○○○○○

Discussion
○○○

Out of vaults

Willing to share?

- Would share
- Would share w/ company, but not public
- No Sharing

Introduction
○○○○○○○○○○○

Open GWAS
○○○○●

Privacy & Implications
○○○○○○○

Discussion
○○○

Out of vaults

Willing to share?

Would you freely publish your genotyping results...

...because of possible personal benefits?	3,64
...because you want to help scientists with their research?	4,53
out of curiosity?	3,5
...because it may deliver advertising that is relevant to you?	1,72

What can happen to your open data?

- Positive and negative consequences

Introduction
0000000000

Open GWAS
00000

Privacy & Implications
●000000

Discussion
000

Some Implications

What can happen to your open data?

- Positive and negative consequences
  - Possibly extremely bad consequences

What can happen to your open data?

- Positive and negative consequences
  - Possibly extremely bad consequences
- Up to you to decide whether you want to open your data

Positive consequences

- More knowledge about yourself

Positive consequences

- More knowledge about yourself
- Cheap, open science

Positive consequences

- More knowledge about yourself
- Cheap, open science
- Great data-source for citizen scientists

Introduction
0000000000

Open GWAS
00000

Privacy & Implications
0000000

Discussion
000

Consequences

Negative consequences

- People know more about you than you might like

Introduction
0000000000

Open GWAS
00000

Privacy & Implications
0●00000

Discussion
000

Consequences

Negative consequences

- People know more about you than you might like
  - Including your boss, insurance company, government...

Negative consequences

- People know more about you than you might like
  - Including your boss, insurance company, government...
- Knowledge isn't static: Future research could show new, negative (or positive) associations.

Negative consequences

- People know more about you than you might like
  - Including your boss, insurance company, government...
- Knowledge isn't static: Future research could show new, negative (or positive) associations.
- Personal SNPs very similar to parents and relatives

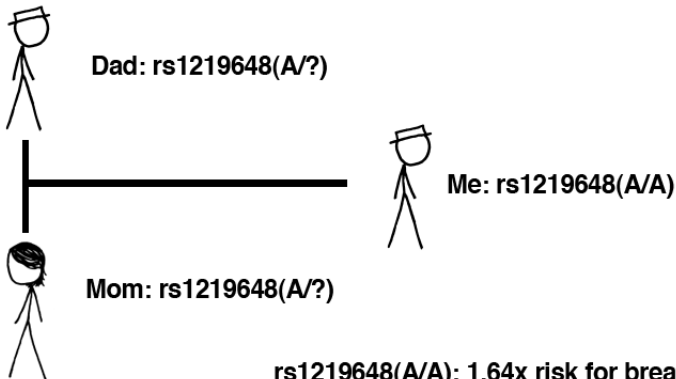# Somebody Else's Problem? A case study

Me: rs1219648(A/A)

rs1219648(A/A): 1.64x risk for breast cancer
rs1219648(A/G): 1.20x risk for breast cancer
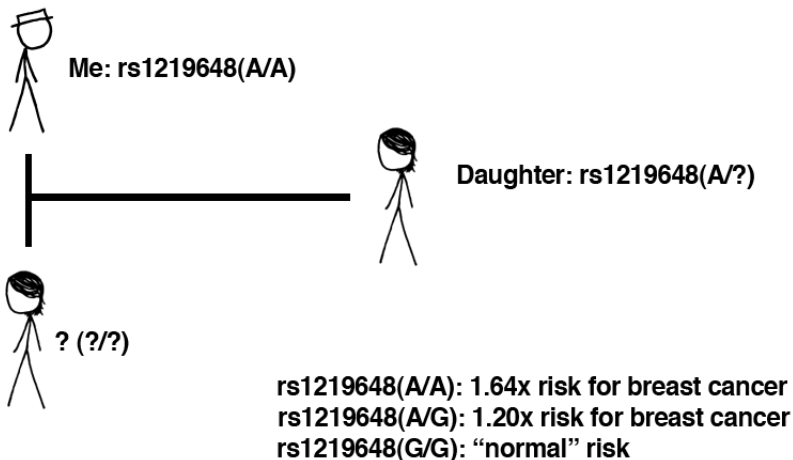rs1219648(G/G): "normal" risk

Somebody Else's Problem? A case study

Dad: rs1219648(A/?)

Me: rs1219648(A/A)

Mom: rs1219648(A/?)

rs1219648(A/A): 1.64x risk for breast cancer
rs1219648(A/G): 1.20x risk for breast cancer
rs1219648(G/G): "normal" risk

Introduction
0000000000

Open GWAS
00000

Privacy & Implications
0000000

Discussion
000

Consequences

Somebody Else's Problem? A case study

Me: rs1219648(A/A)

Daughter: rs1219648(A/?)

? (?/?)

rs1219648(A/A): 1.64x risk for breast cancer
rs1219648(A/G): 1.20x risk for breast cancer
rs1219648(G/G): "normal" risk

Possible Solutions

- What about laws?

Possible Solutions

- What about laws?
    - US: Genetic Information Nondiscrimination Act (GINA, 2008)

Possible Solutions

- What about laws?
  - US: Genetic Information Nondiscrimination Act (GINA, 2008)
  - Germany: Gendiagnostikgesetz (GenDG, 2010)

Introduction
○○○○○○○○○

Open GWAS
○○○○○

Privacy & Implications
○○○○○○○

Discussion
○○○

Open GWAS

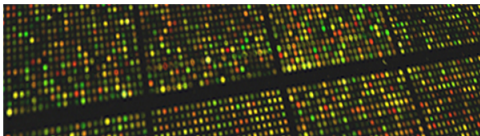openSNP   News   Genotypes   Phenotypes   SNPs   Users   [Search here]   Sign in   FAQ

Welcome to *openSNP*

openSNP allows customers of direct-to-customer genetic tests to publish their test results, find others with similar genetic variations, learn more about their variations, find the latest primary literature on their variations and help scientists to find new associations.

[Sign Up!]

For Genotyping Users | For Scientists | FAQ

Upload Your Genotyping File

Upload the genotyping raw-data you got from 23andMe or deCODEme to the database of *openSNP* to make it available for everybody.

Share Your Phenotypes & Traits

Share as many phenotypes, characteristics and traits with other *openSNP* users and find others with similar characteristics. And maybe help scientists to discover new genetic associations.

Share your stories on variations & phenotypes

openSNP lets you share your stories on your genetic variations & phenotypes with others. Discover the stories of other users. Find others to exchange experiences about your variations.

Find literature on genetic variation

openSNP gets the latest open access journal articles on genetic variations via the Public Library of Science. Additionally popular articles are indexed via the social reference manager Mendeley. Summaries are provided by SNPedia.

Introduction
0000000000

Open GWAS
00000

Privacy & Implications
0000000

Discussion
000

openSNP

- No central repository for open genotypings!

openSNP

- No central repository for open genotypings!
- We've created openSNP.org

openSNP

- No central repository for open genotypings!
- We've created openSNP.org
- open source repository for CC0-genotypings from 23andme, deCODEme and others

... continued

- Allows users to annotate with phenotypes (hair colour, nicotine dependence...)

... continued

- Allows users to annotate with phenotypes (hair colour, nicotine dependence...)
- Everybody can download everything

... continued

- Allows users to annotate with phenotypes (hair colour, nicotine dependence...)
- Everybody can download everything
- So far: 78 genotypings and 188 users

Conclusions

- Open GWAS are the future of personalised medicine

Conclusions

- Open GWAS are the future of personalised medicine
- It's in the hands of users to make or break the situation

Conclusions

- Open GWAS are the future of personalised medicine
- It's in the hands of users to make or break the situation
- Chance to take science into our own hands

Introduction
0000000000

Open GWAS
00000

Privacy & Implications
0000000

Discussion
●○○

Outlook

Future of openSNP

- We've won the PLoS/Mendeley Binary Battle

Introduction
0000000000

Open GWAS
00000

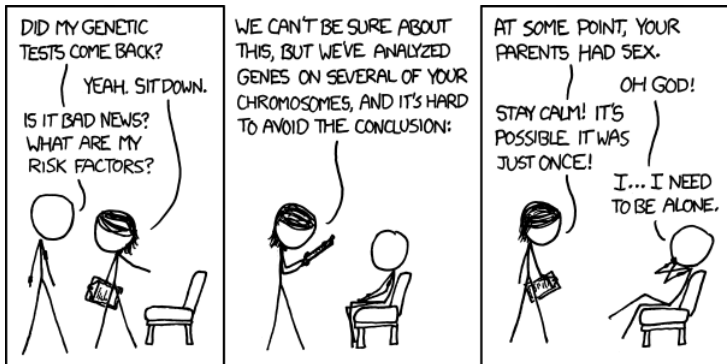Privacy & Implications
0000000

Discussion
●00

Outlook

Future of openSNP

- We've won the PLoS/Mendeley Binary Battle
- Constantly improving the project

Future of openSNP

- We've won the PLoS/Mendeley Binary Battle
- Constantly improving the project
- Trying to get funds for free genotypings

Introduction
0000000000

Open GWAS
00000

Privacy & Implications
0000000

Discussion
0●0

Outlook

The end

Thanks for listening. Any questions?

References

Do *et al.* (2011) Web-Based Genome-Wide Association Study Identifies Two Novel Loci and a Substantial Genetic Component for Parkinson's Disease. PLoS Genetics 7(6): e1002141. doi:10.1371/journal.pgen.1002141

Eriksson *et al.* (2010) Web-Based, Participant-Driven Studies Yield Novel Genetic Associations for Common Traits. PLoS Genet 6(6): e1000993. doi:10.1371/journal.pgen.1000993

Kogan, *et al.* (2011): Thin-slicing study of the oxytocin receptor (OXTR) gene and the evaluation and expression of the prosocial disposition. Proceedings of the National Academy of Sciences

Sladek *et al.* (2007): A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 445 (7130): 881-5.

The Wellcome Trust Case Control Consortium (2007): Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447: 661-678.