DNA digest

open access and secure data digest

Open Data on the Web
Fiona Nielsen, April 24, 2013

DNA digest

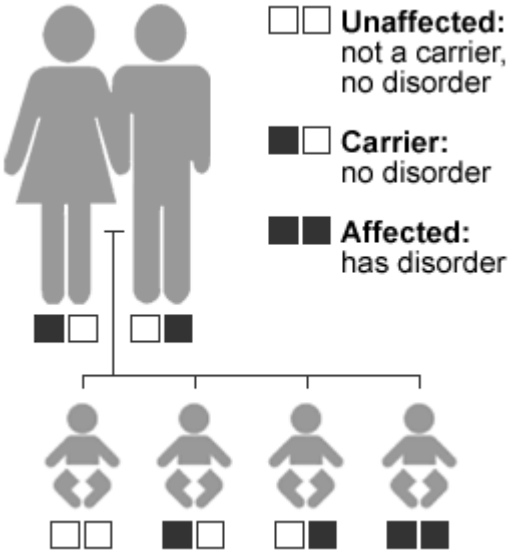
DNAdigest enables data access

- DNA sequencing
- Genome research today
- Why genomics data is not open
- How DNAdigest will enable data access

DNA digest

DNA sequencing

- cancer research
- heritable traits and illness
- rare diseases



images :http://www.pbs.org/newshour/rundown/2012/10/new-rapid-test-can-diagnose-babys-genetic-disorder-in-days.html
and http://www.nidcd.nih.gov/health/hearing/pages/usher.aspx

DNA digest

New technology, new opportunities

The technology is developing fast, promising a **Genomic Revolution**,

DNA digest

New technology, new opportunities

The technology is developing fast, promising a **Genomic Revolution**,

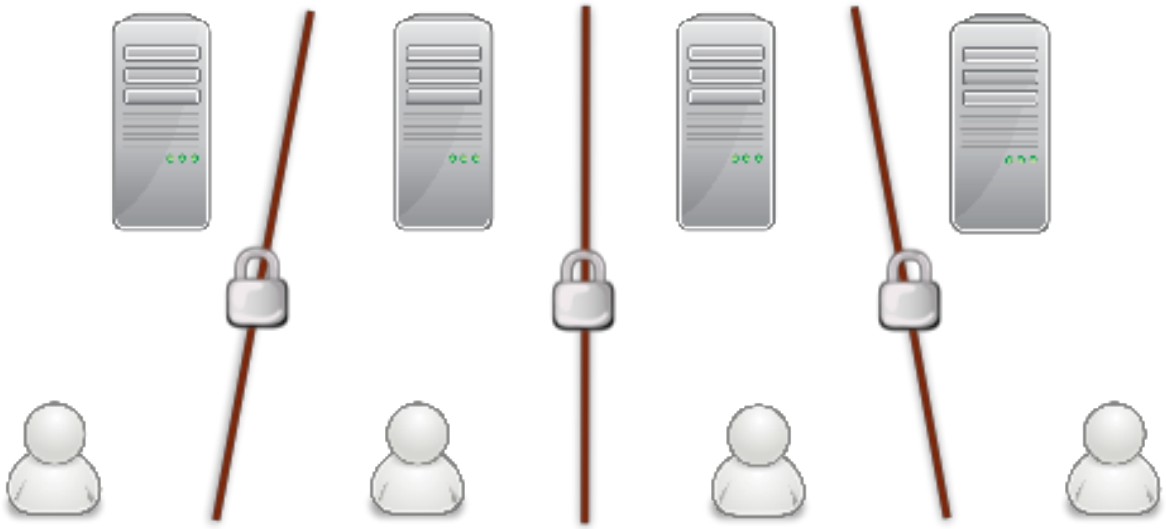
But the research progress is not as fast as it could be...

DNA digest

New technology, new opportunities

Why is research lagging behind?

- Because researchers are not sharing all their data!



DNA digest

Genome research today

the patient

the researcher

the data

DNA digest

Genome interpretation is hard

To make sense of his data, the researcher needs more samples for comparison

- He will look in databases of genomic variation
- He will search in literature describing the disease

But still the common case is a large number of ***variants of unknown effect***

DNA digest

More data needed to validate results

At the same time, other researchers around the World are facing the same problem:

- They want to learn from other data sets
- But they do not make their own data openly accessible

DNA digest

Why is genomics data not shared?

Data sharing is a problem:

- medical data confidential
- ability to identify individuals from genome data
- bulky data sets

DNA digest

Current practice for sharing

Patients give consent for their data to be published and shared only if:

- de-identified
- aggregated in research result

Current practise is to publish results in scientific journals, **no open sharing of raw data**

DNA digest

How to open up access to the data?

DNA digest

How to open up access to the data?

Solution:

Not all research questions require access to entire data sets

DNA digest

How to open up access to the data?

Solution:

Not all research questions require access to entire data sets

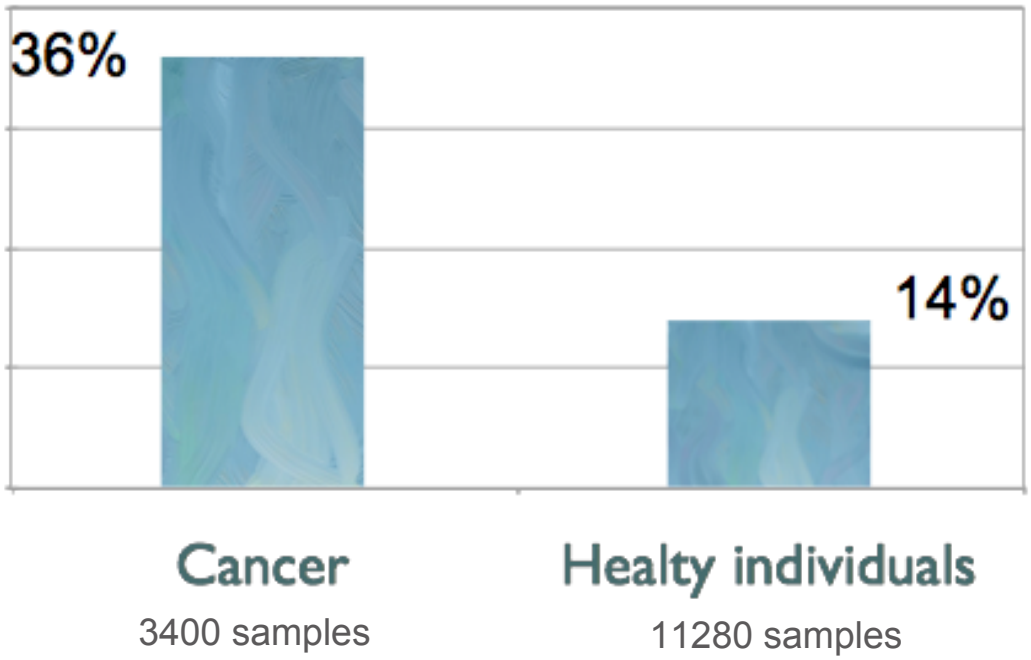
Example:

Does this mutation occur at higher frequency in cancer than in healthy samples?

DNA digest

Example Use Case

"What is the frequency of this mutation in disease samples versus reference samples?"



DNA digest

DNAdigest allows data access

Advantages:

- Anonymisation by aggregation
- Access can be made open for all users
- Queries for direct hypothesis testing

Challenges:

- Connect existing repositories with the DNAdigest API
- Create incentives for institutions to connect
- Build sustainable business model

DNA digest

DNAdigest needs your help

We propose a solution, but we need your help to make it happen:

- Inform patients that their data is not shared sufficiently today
- Encourage researchers to share their data
- Support incentives that promote data sharing

DNA digest

Summary

DNAdigest will allow interactive queries, collecting data from multiple repositories with results presented aggregated and anonymised

DNAdigest is a not-for-profit organisation, founded for the purpose of enabling secure open access sharing of genomic data.

Please visit us at DNAdigest.org

Thank you!