

Temporal dynamics of collaborative networks driven by large scientific consortia:

Mining an ENCODE “Data Exhaust”

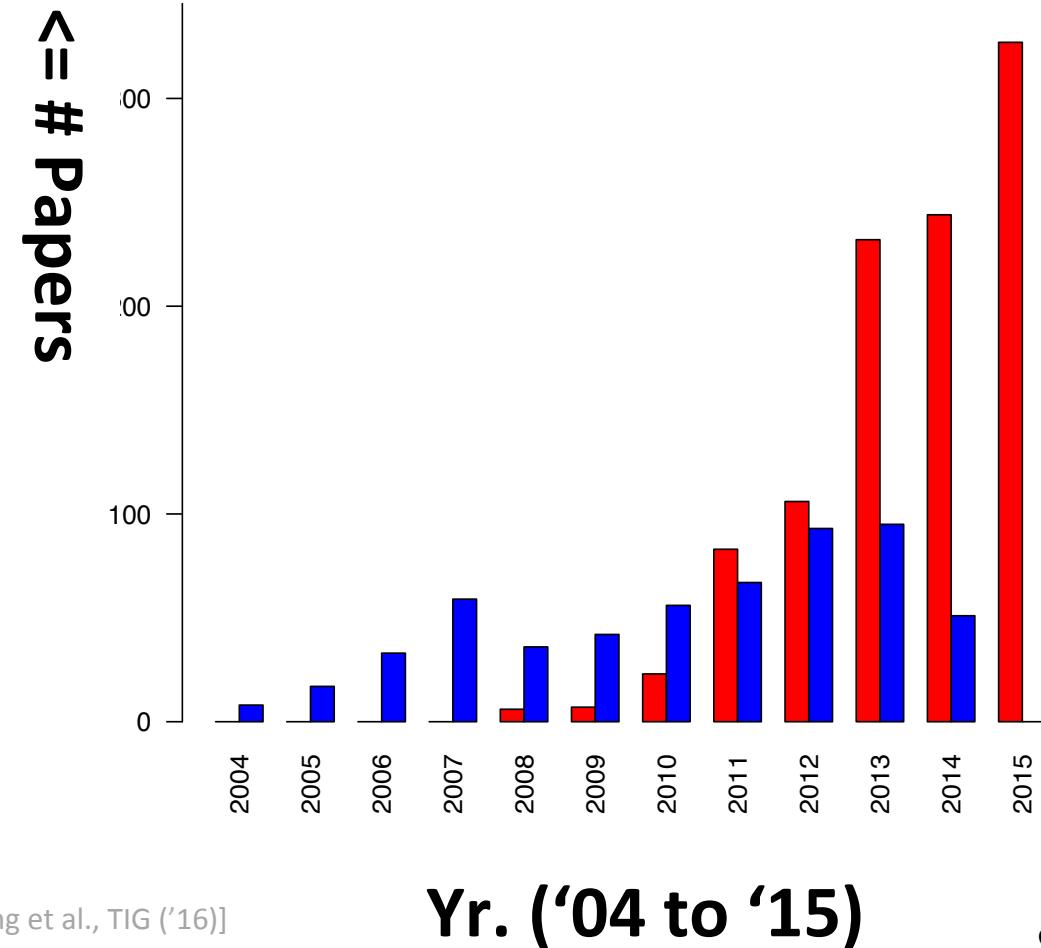
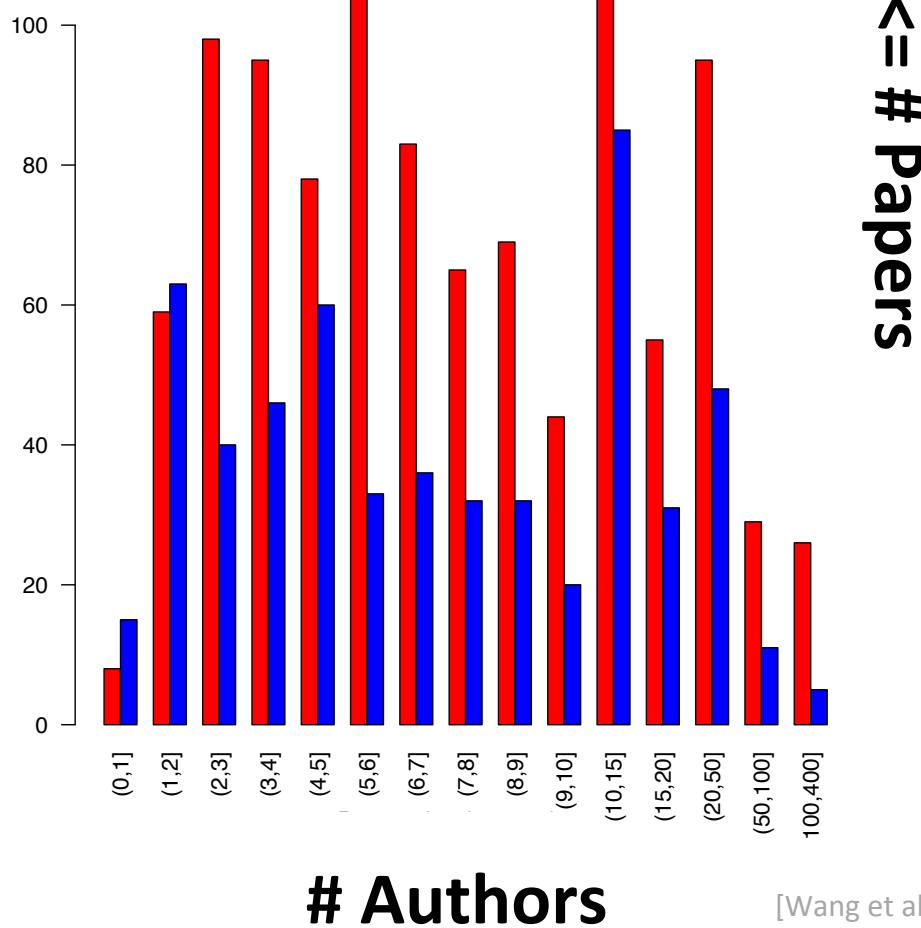
D Wang

KK Yan, J Rozowsky, E Pan

M Gerstein

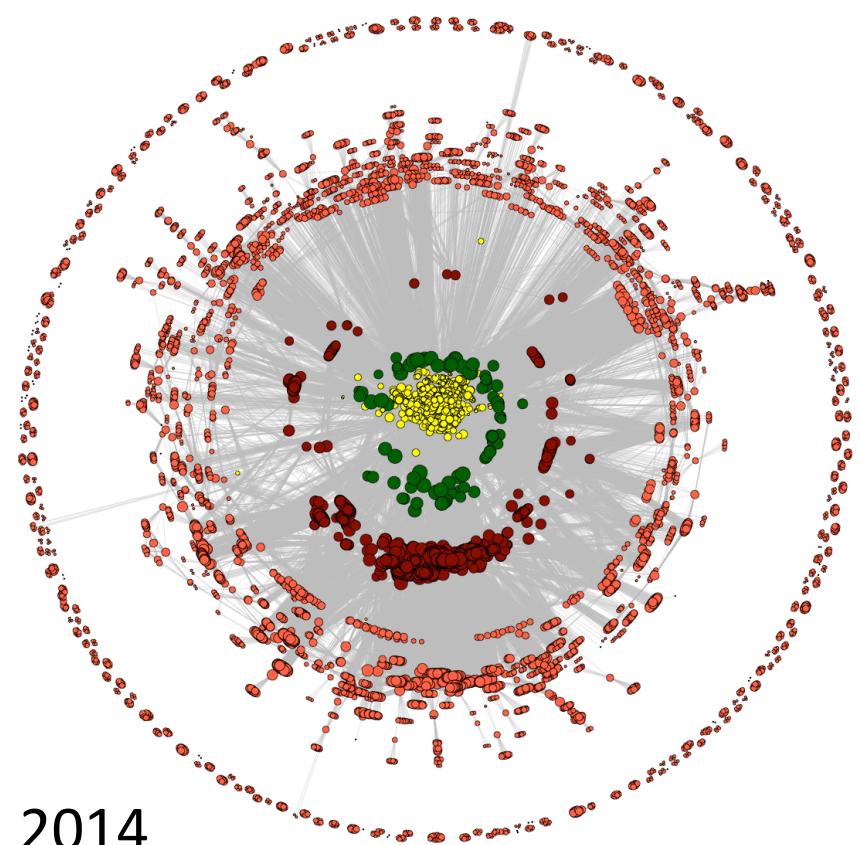
With help of M Pazin at NHGRI, identified: **702 community papers that used ENCODE data but were not supported** by ENCODE funding & **558 consortium papers supported by ENCODE funding** (<https://www.encodeproject.org/search/?type=Publication> for up-to-date query) Then identified **1,786 ENCODE members** & **8,263 non-members**.

■ non-ENCODE (papers used ENCODE data) ■ ENCODE



Co-authorship Network of ENCODE members & Data Users

- ENCODE member
- non-member
- ENCODE member broker
- non-member broker
- co-authorship

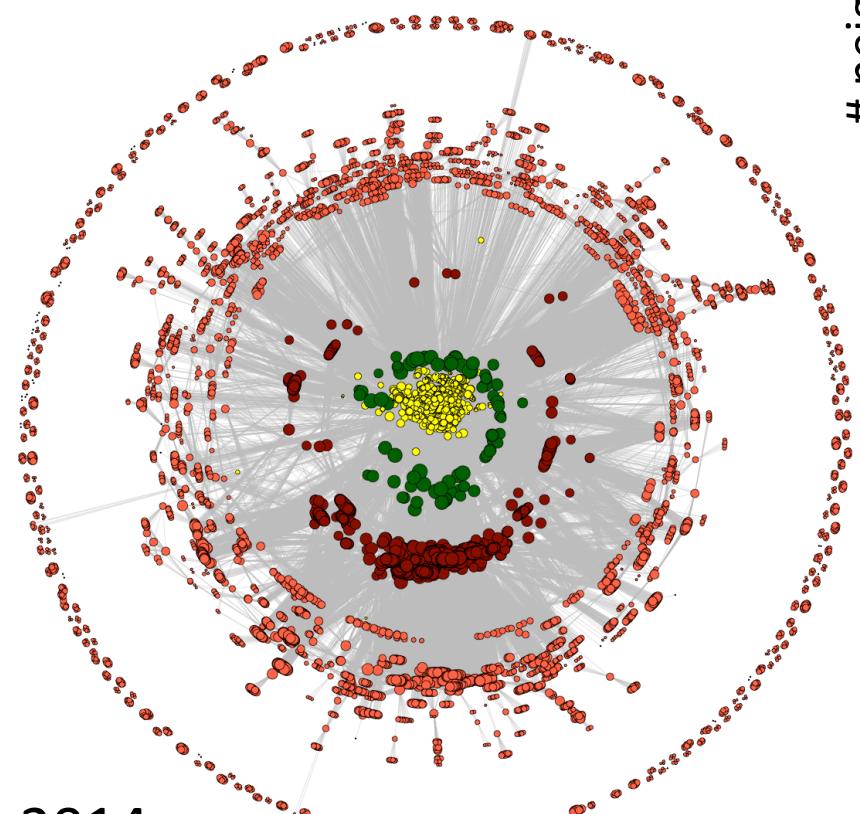


2014

[Wang et al., TIG ('16)]

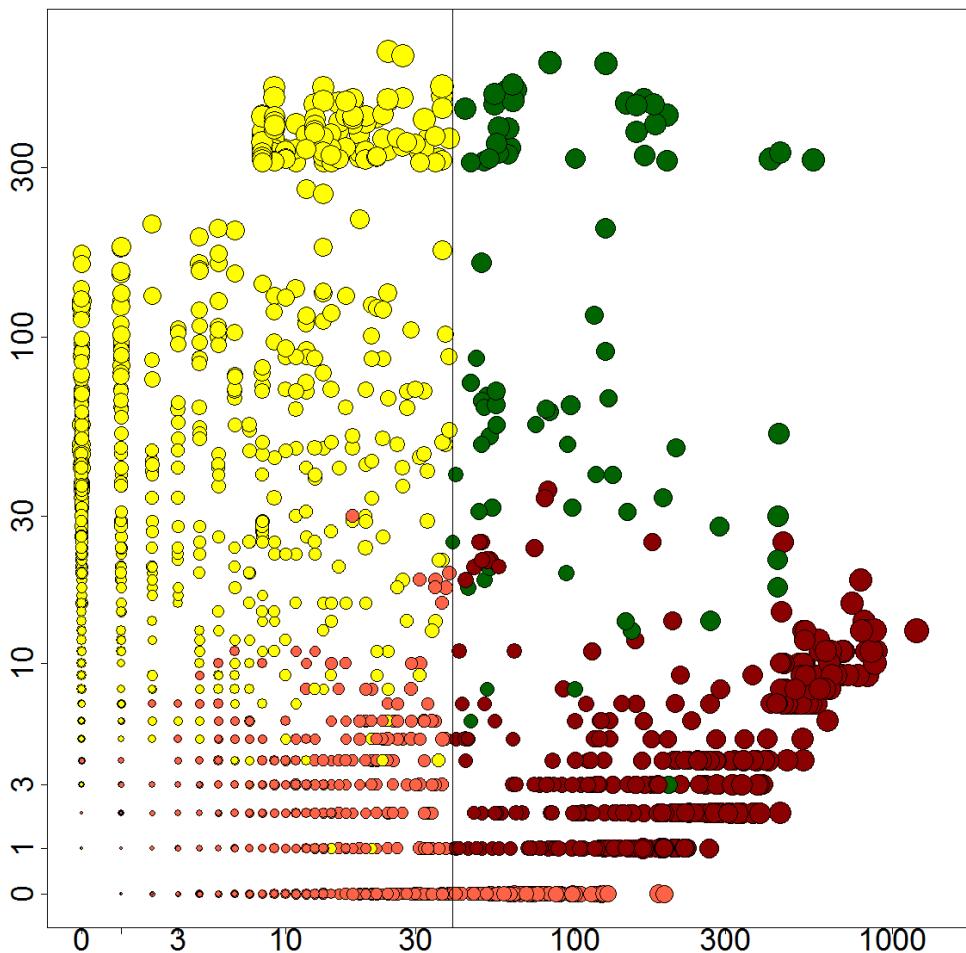
Co-authorship Network of ENCODE members & Data Users

- ENCODE member
 - non-member
 - ENCODE member broker
 - non-member broker
- co-authorship



2014

neighbors: non-ENCODE ==>

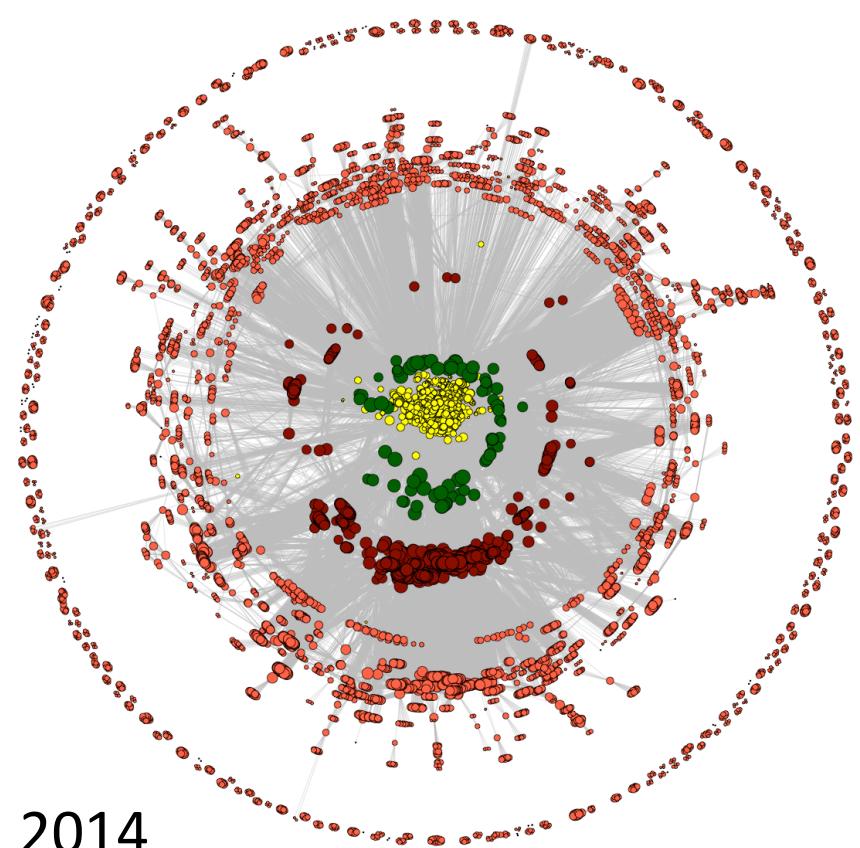


neighbors: ENCODE ==>

[Wang et al., TIG ('16)]

Co-authorship Network of ENCODE members & Data Users

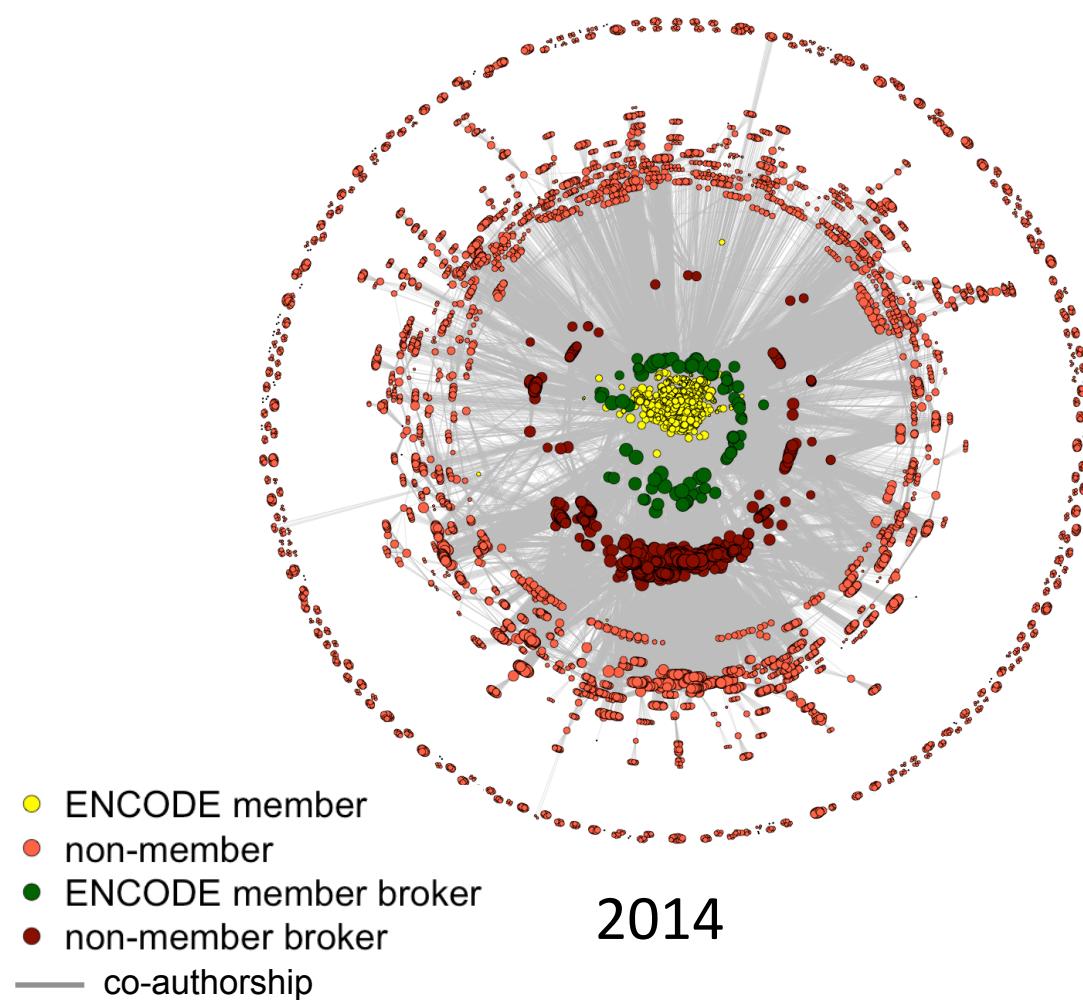
- ENCODE member
- non-member
- ENCODE member broker
- non-member broker
- co-authorship



2014

[Wang et al., TIG ('16)]

Dynamics of co-authorship network



Dynamics of co-authorship network

2009

2010

2008

2007

2006

2005

2004

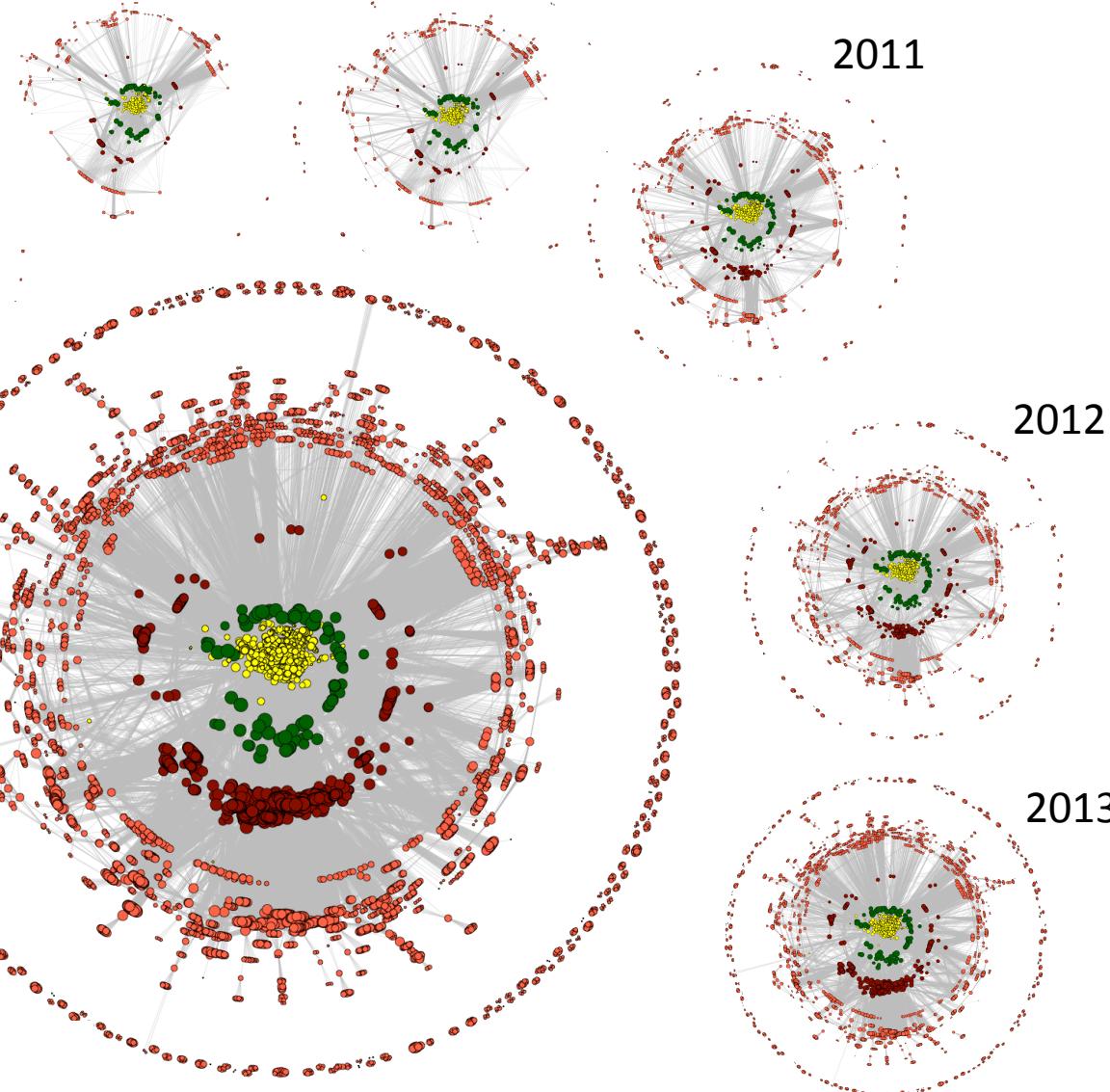
2014

2011

2012

2013

- ENCODE member
- non-member
- ENCODE member broker
- non-member broker
- co-authorship



Dynamics of co-authorship network

2009

2010

2008

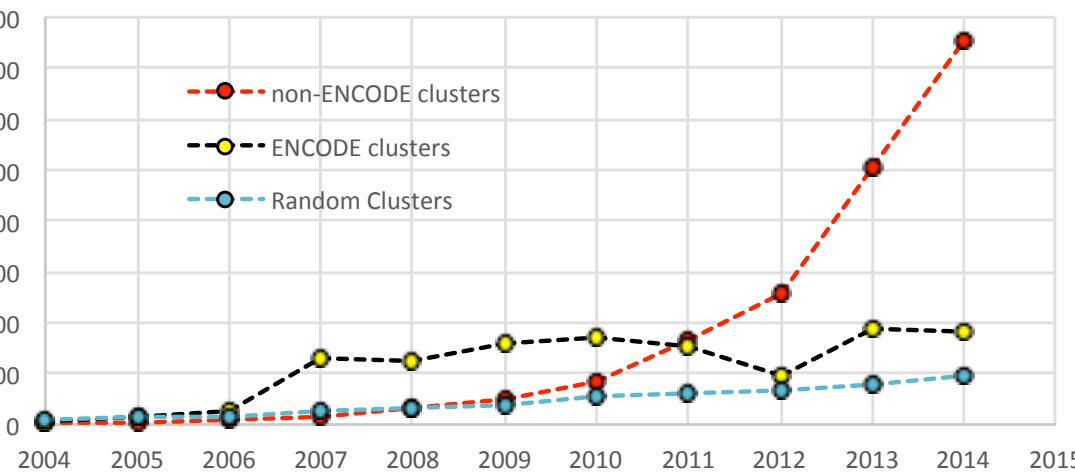
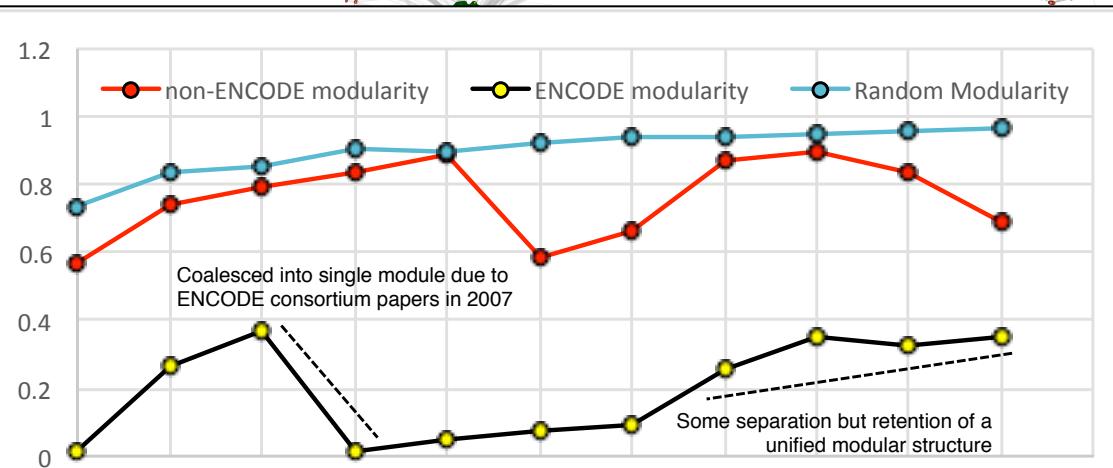
2011

2012

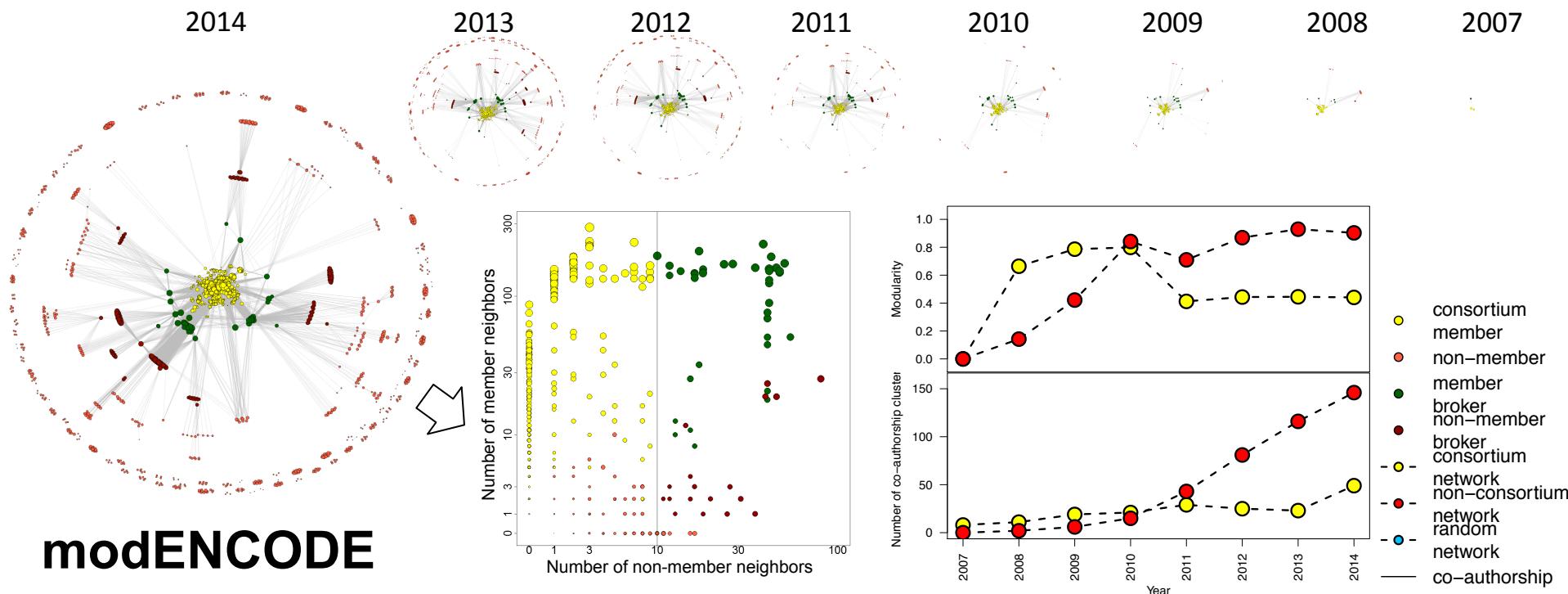
2013

“Modularity”

Number of clusters



Similar findings in terms of slow growth trends & broker scientists in the modENCODE consortium as for ENCODE



- **Conclusions**
 - Value of publication patterns generated by the consortium
 - Co-authorship network statistics relate to publication rollouts & show **gradual adoption by a diverse** community
 - Key role of **brokers** in data dissemination