1

Mapping the Faulty Process in Breakthrough Misses

Sen Chai - ESSEC

Organization Science Winter Conference

February 5, 2016

Motivation

- Scientific & technological breakthroughs spur wealth creation and economic growth (Rosenberg, 1974)
- Lots written on breakthrough successes
- Much less on failures to breakthrough
- Failure mechanisms may be qualitatively different from success of breakthrough (Hackman, 2009)
- Help researchers & managers of creative firms
  - Mitigate delays and avoid misses
  - Better use of resources to boost performance

WHY ARE BREAKTHROUGHS MISSED?

Some definitions

- Breakthroughs
  - Advances that disturb the previous understanding of phenomenon in a fundamental manner by irreversibly altering the paradigm upon which that understanding rests (Kuhn, 1962)
- Paradigm
  - View, model, or pattern commonly accepted and shared by a research community (Kuhn, 1962)

Ongoing debate about sources of breakthrough

- Debate in creativity literature (Weisberg, 1999)
  - Foundation view – deep immersion in a particular domain is required for creative work
    - basic elements upon which new ideas are constructed (Gardner, 1993)
    - to notice anomalies and potential gaps for breakthrough discoveries (Kaplan & Vakili, 2015; Taylor & Greve, 2006)
  - Tension view – too much depth of knowledge may be damaging
    - constrains idea sources to a narrow domain (Simonton, 1999)
    - limits the recombination of distant and diverse knowledge (Hargadon & Sutton, 1997)

Uncertainty about sources of breakthrough

- Productivity
  - More productive (Simonton, 1999) vs. Less productive
  - Highly cited
  - Prestige
- Connectivity
  - Collaborative (Singh & Fleming, 2010) vs. Lone scientist vs. Working alone during certain stages (Girotra et al., 2010)
  - Brokers (Burt, 2004) vs. Cohesive position (Obstfeld, 2005; Uzzi, 1997)
  - Core (Collins, 1998; Gieryn & Hirsh, 1983) vs. Periphery (Jeppesen & Lakhani, 2010)
  - Mobility (McEvily & Zaheer, 1999)
- Personal Attributes
  - Specialist vs. Generalist (Dougherty, 1992; Leonard-Barton & Swap, 1999)
  - More experience (Jones, 2009) vs. Less experience (Simonton, 1989)
  - Academic (Stern, 2004) vs. Industry Affiliated

Reconcile with view of discovery as process

- Both depth and breadth of knowledge are needed
  - Depth to see anomalies and identify openings to potential breakthrough (Kaplan & Vakili, 2015; Taylor & Greve, 2006)
  - Breadth to solve the breakthrough problem once the anomaly has been identified (Hargadon & Sutton, 1997)

- Process by which scientific discoveries and technological inventions emerge
  - Paradigm shifts (Kuhn, 1962)
    - Accumulated anomalies & crisis
    - Extraordinary science & resolution

Why are breakthroughs missed?

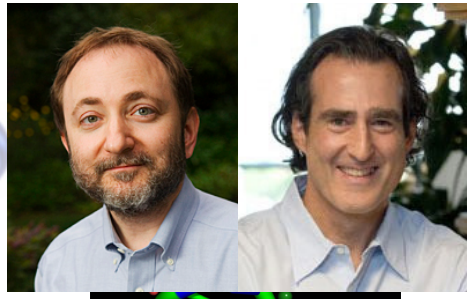
- Breakthrough as process of multiple failures with eventual success

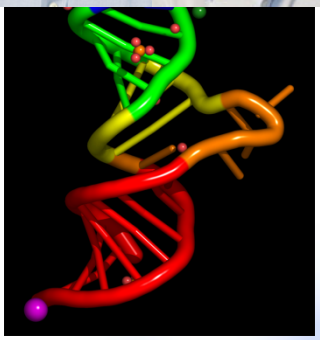
- Counterfactual of why breakthroughs are missed or delayed, by interviewing scientists on verge of breakthrough
- Failures at different stages of process
- Qualitative case study to capture cognitive process

Setting

RNA interference (RNAi)

- Natural mechanism that silences genes by turning them on and off
- Challenges central dogma
- Trigger (dsRNA) found in 1998
- Nobel prize awarded in 2006





New drug class

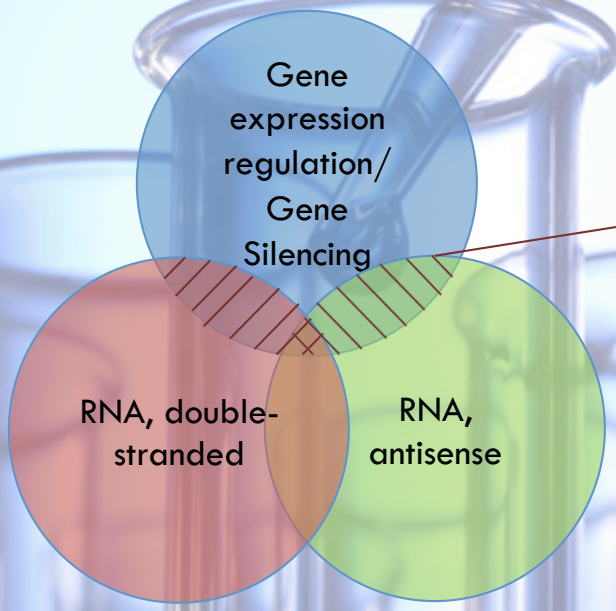
Interviewee identification

- Challenge identifying
  - community of researchers *before* breakthrough
  - scientists on verge of breakthrough

- Community – scientists who share the same research interests
  - Functional definition - Write publications in similar topics
    - Noisy sample, omit scientists who shifted their research
  - Conference definition - Attend the same conferences
    - Omit non-attendees

Using bibliometric data for functional definition

PubMed Author-ity (Torvik & Smalheiser, 2009)



1,551 papers published until 1999
3,959 unique authors
144 newcomers

Using bibliometric data for functional definition

- Researchers close to breakthrough *in* 1998, but missed it
  - 1998 publications in top 1% of forward citations – 6 interviewees (9 requests)
- Researchers close to breakthrough *before* 1998, predicted to do well in 1998 but missed it
  - 1998 publications with high residuals of forward citations – 7 interviewees (10 requests)
    - Outcome variable: Author's aggregate forward citation count for publications *during* year of discovery
    - Explanatory variables: Sources of creative breakthroughs *prior* to discovery

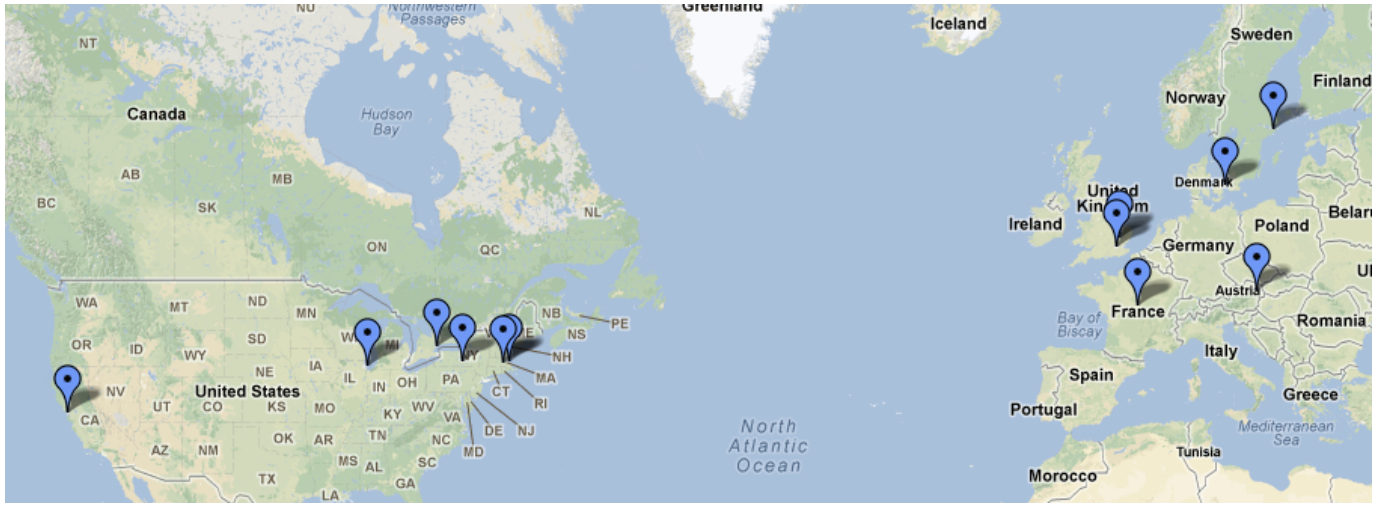
Using conferences for community definition

- Presenters of the earliest RNA silencing conference
  - Prominent enough to present a paper at the conference
  - 5 attendees overlap with interviewees from functional definitions
  - 5 additional interviewees (8 requests)

- Snowball sampling to substantiate
  - Ask respondents who of their colleagues to interview

Interviews

- 18 interviewees (27 requests)
  - 1 Nobel Prize winner Physiology and Medicine
  - 3 Lasker Award winners for Basic Research in Medicine

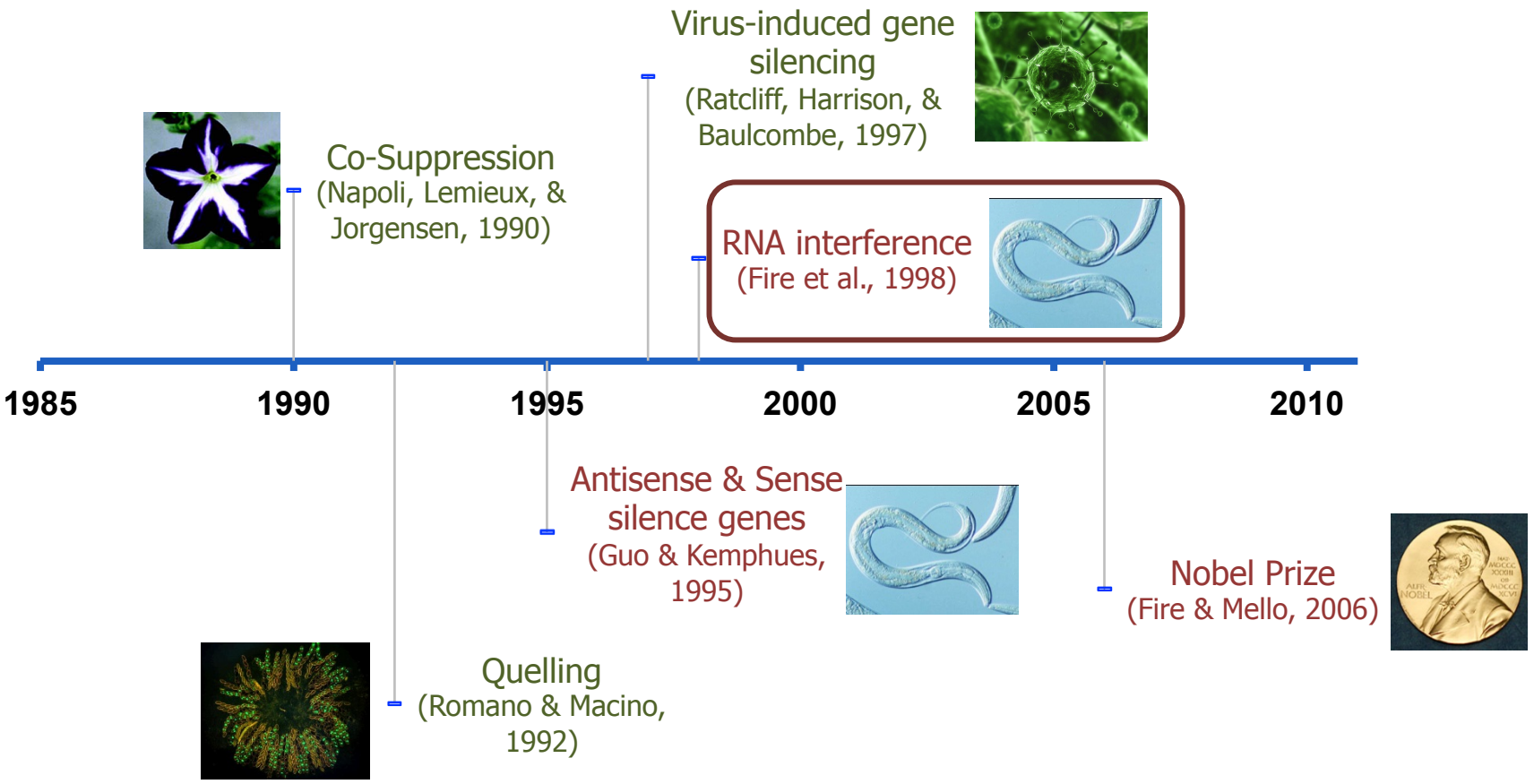


- Semi-structured
  - Breakthrough and circumstances around it
  - Scientist's productivity and focus
  - Community definition and characteristics

Archival data

- Mitigate hindsight bias (Golden, 1992)
- Triangulate
  - quotes extracted from RNAi-related research articles of each interviewees
  - citation patterns of articles

Multiple observations before flash of understanding



Finding

- Seminal discovery missed several times in different communities

- Not noticing or recognizing anomalies
- Actively resisting solutions
- Not communicating across communities

Not noticing or recognizing anomalies

Animal community was blinded by pursuit of normal science

- Preoccupied with experiment at hand

"Once we knew it was a gene specific effect, we didn't really care how it worked. All we cared about was that we could ***use it to advance our initial experiments.****"*

- Used phenomenon as tool/technique instead of seeing it as topic worthy of inquiry

"When people tried it and it worked, it was like ok let's work with it. ***Very few people thought it was worth studying. But everybody wanted to use it.****"*

- Paradigmatic rigidity
  - Prevented abductive reasoning: no doubt of anomaly

Actively resisting solutions

Thinking constrained by paradigm – central dogma of molecular biology

- Animal community did not recognize anomaly still could have improved the technique, but

*"It's weird and not expected because **we all knew that if we make dsRNA that's a dead end**. It's inhibition of the other RNA, you can't use that to make something."*

- Plant community did recognize anomaly, but

*"Why didn't the plant people get to where Fire & Mello did? My main insight is that **we were so focused on transgenes to manipulate DNA expression**. We never got to introducing RNA. It was regarded as unstable – that was never going to work."*

- Paradigmatic rigidity
  - Constrained search to local search
  - Prevented abductive reasoning: no proposal of significant novel solution

Not communicating across communities

Unable to Connect the Dots

- Disconnect between sub-communities

*"The paper […] on petunias was published in […] the journal Plant Cell. Really **nobody followed that**. Nobody!" (animal scientist)*

*"During the initial years when we were working on [understanding gene silencing] **we weren't talking with animal people very much**." (plant scientist)*

- Belief that each community in its own separate paradigm – anomalies are unique to their own organism of study, not evolutionarily conserved

*"As far as we knew at the time, it was a **very specific phenomenon for c. elegans**." (animal scientist)*

*"And at the time **I don't think we were thinking too much about necessarily the animal work**." (plant scientist)*

Not communicating across communities

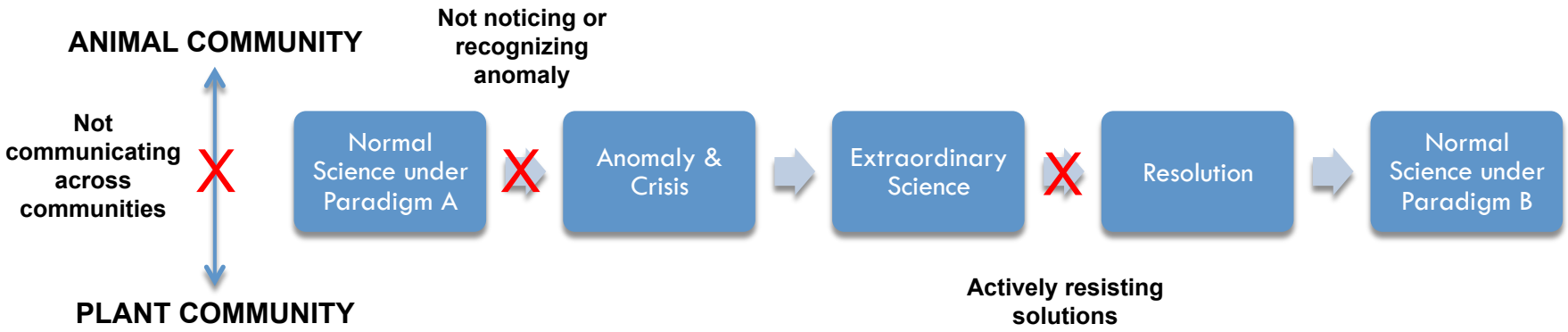
- Thought of being in separate paradigms
  - Constrained search to local search within community
  - Prevented abductive reasoning: no use of solution from other organisms

- Each community had one piece of the puzzle
  - Animal community: single-stranded **RNA**
  - Plant community: **double-stranded** DNA

⇨ Trigger: *dsRNA*

Summary of failures

- Cognitive bias
  - Not noticing or recognizing anomalies
  - Actively resisting solutions
- Structural barrier that led to cognitive bias
  - Not communicating across communities



Implications to research

- Nuanced interpretation for the debate between depth and breadth of knowledge
- Evidence in line with the tension view
  - Normal science – fail to notice anomaly
  - Extraordinary science – active resistance of solution
- Paradigms are filters
  - constrain and determine the relevance of knowledge produced
  - hinder creativity especially with radical outcome

Implications to practice

- Avoid the pitfalls identified and potentially intervene

- Expand research scope & explore at the fringe
- Exploit social ties for substantiation
- Broaden exposure and awareness of work
  - Facilitate cross-boundary research teams
  - Establish external collaborative links through seminars and conferences
- Incentivize for broader intellectual reach

Future Research

- Generalization
  - Breakthrough datasets
  - Qualitative analysis
    - Illuminate mechanism and processes in production of knowledge
    - Develop hypotheses for quantitative tests
  - Quantitative analysis
    - Test failures using bibliometrics
    - Content analysis

Thank You!

Supplemental materials

Not noticing or recognizing failures

Confirmatory analysis in interviewee's publications

- Antisense RNA consistently mentioned as a technique

  "Progress in identifying additional components could occur quite rapidly because of the recent development of a simple ***application of antisense RNA technology to c. elegans.*** *[...]* ***Using this technique****, candidate genes identified by a variety of means can be assayed rapidly for important roles in cell polarity." (Guo & Kemphues, 1995)*

- Antisense RNA included in the publication's Methods section as a tool

Actively resisting solutions

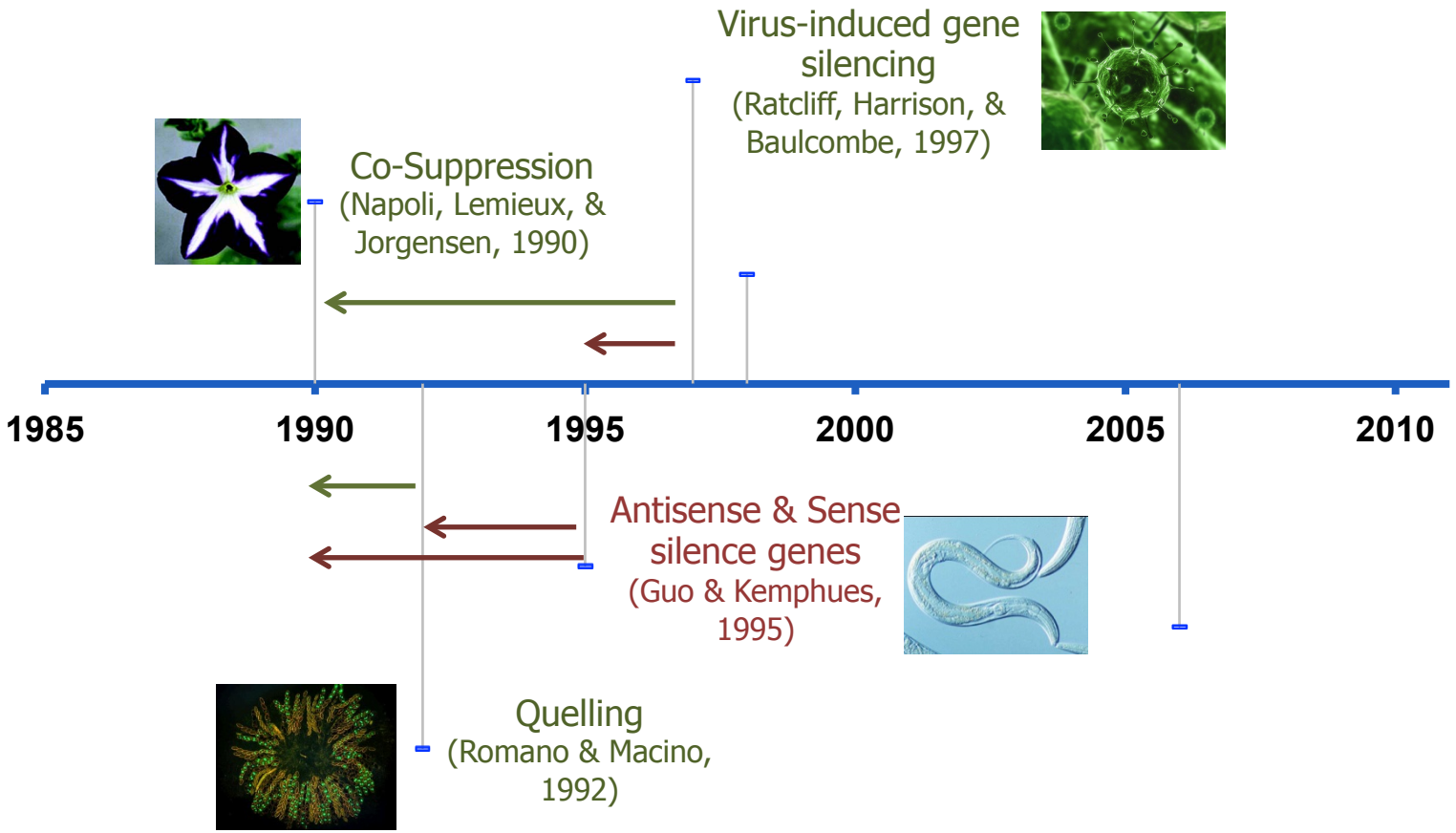
Confirmatory analysis in interviewee's publications

- Animal community: only antisense RNA mentioned as tool, no mention of sense RNA or dsRNA
- Plant community: only DNA-based transgenes and DNA-methylation mentioned

"Unexpectedly, the introduced gene created a block in anthocyanin biosynthesis. [...] The mechanism responsible for the reversible co-suppression of homologous genes in trans is unclear, but the erratic and reversible nature of this phenomenon suggests the possible involvement of ***methylation****." (Napoli et al., 1990)*

Not communicating across communities

- Confirmatory analysis find distinct citation patterns between plant/fungus and worm communities



What did Fire and Mello do differently?

- First contact with phenomenon as tool
- Noticed puzzle and realized that the phenomenon itself was interesting, important and worth studying

*"Andy and I knew that RNA interference was **something incredible** when we started working with it."*

- Unconstrained by current paradigm

*"[…] the concept that **double-stranded RNA** might be a component of the injected material was **hardly a leap of logic.**"*

- Nobel paper cited articles in plants and c. elegans

- Not marked by setbacks/failures
- Failures not apparent if only taken success perspective

Predicting RNAi discovery

Theory	**Explanatory Variable (data prior to 1998)**
Publication history	# prior publications
Publication eminence	# forward citations
Collaboration vs. lone scientist	# co-authors
Brokerage vs. cohesion	Burt's constraint
Core vs. periphery	technical & collaborative core
Generalist vs. specialist	publication depth
Lifecycle	experience
Affiliation type	academic
Prestige	prestige & prestige dummy
Mobility	newcomer & multiple affiliations

Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
nobeldum	3959	0.002	0.039	0.000	1.000
top10cite	3959	0.100	0.300	0.000	1.000
lncite98	3959	1.761	1.756	0.000	8.009
lnpub98	3959	0.888	0.822	0.000	4.477
lnpub97	3959	2.585	1.371	0.000	7.073
lncite97	3959	4.178	2.058	0.000	9.692
lnpub97_fl	3959	1.934	1.426	0.000	6.836
lncite97_fl	3959	3.193	2.300	0.000	9.533
constraint	3959	0.657	0.315	0.000	1.932
lncoauthor	3959	2.279	1.150	0.000	5.855
collabcore	3959	0.821	0.383	0.000	1.000
techcore	3959	0.450	0.498	0.000	1.000
pubdepth	3959	0.281	0.058	0.000	1.000
lexp	3959	2.354	0.883	0.000	4.174
academic	3959	0.996	0.044	0.000	1.000
prestige	3910	5.889	10.683	0.000	50.000
prestiged	3910	0.347	0.476	0.000	1.000
newcomer	3959	0.036	0.187	0.000	1.000
affil1p	3910	0.591	0.492	0.000	1.000

Correlation matrixes

	nobeldum	top10cite	lncite98	lnpub98
nobeldum	1			
top10cite	0.1167	1		
lncite98	0.1169	0.6044	1	
lnpub98	0.0168	0.4711	0.815	1

	lnpub97	lncite97	lnpub97_fl	lncite97_fl	constraint	lncoauthor	collabcore	techcore	pubdepth	lexp	academic	prestige	prestiged	newcomer	affil1p
lnpub97	1.000														
lncite97	0.815	1.000													
lnpub97_fl	0.939	0.771	1.000												
lncite97_fl	0.788	0.888	0.861	1.000											
constraint	-0.210	-0.225	-0.202	-0.216	1.000										
lncoauthor	0.649	0.568	0.529	0.455	-0.089	1.000									
collabcore	0.214	0.325	0.151	0.222	-0.361	0.263	1.000								
techcore	-0.219	-0.076	-0.218	-0.104	0.033	0.017	0.063	1.000							
pubdepth	0.318	0.165	0.302	0.202	-0.069	0.178	-0.005	-0.122	1.000						
lexp	0.847	0.724	0.784	0.674	-0.158	0.551	0.252	-0.120	0.234	1.000					
academic	-0.004	0.020	0.017	0.005	0.001	-0.003	0.008	0.011	-0.001	-0.006	1.000				
prestige	-0.058	0.094	-0.044	0.065	-0.014	-0.059	0.066	0.084	-0.063	-0.048	-0.027	1.000			
prestiged	0.100	0.231	0.110	0.206	-0.062	0.016	0.087	0.024	0.000	0.108	-0.019	0.755	1.000		
newcomer	-0.371	-0.401	-0.268	-0.274	-0.001	-0.392	-0.420	-0.178	0.001	-0.519	0.017	-0.108	-0.143	1.000	
affil1p	0.677	0.599	0.627	0.568	-0.173	0.394	0.166	-0.165	0.188	0.698	0.002	-0.056	0.122	-0.235	1.000

Regression results – Robustness checks

DV	Relogit Nobel nobeldum b/se	Logit Top10c top10cite b/se	QML impact ncite98 b/se	QML prod npub98 b/se
lnpub97_fl	4.553	0.365**	0.217+	1.047**
	(3.120)	(0.120)	(0.120)	(0.050)
lncite97_fl	-2.784	0.863**	0.716**	-0.021
	(1.810)	(0.070)	(0.050)	(0.020)
lncoauthor	-3.063*	0.199*	-0.045	0.004
	(1.210)	(0.080)	(0.060)	(0.030)
constraint	16.731+	-1.064**	-0.583*	-0.257**
	(9.180)	(0.290)	(0.240)	(0.100)
collabcore		0.107	0.190	0.031
		(0.270)	(0.160)	(0.100)
techcore		-0.276*	0.024	0.027
		(0.140)	(0.110)	(0.050)
pubdepth	11.711	0.555	0.600	1.000*
	(17.390)	(1.430)	(1.120)	(0.500)
lexp	0.912	-1.565**	-0.987**	-0.885**
	(2.490)	(0.200)	(0.160)	(0.070)
prestige	-0.002	0.013+	0.008	0.003
	(0.080)	(0.010)	(0.010)	0.000
prestiged		-0.011	-0.083	-0.149*
		(0.170)	(0.110)	(0.060)
academic		0.893	1.624	1.429*
		(1.980)	(1.000)	(0.720)
newcomer		1.104+	0.852*	-0.354+
		(0.590)	(0.390)	(0.180)
affil1p		0.518*	0.174	0.181+
		(0.240)	(0.200)	(0.100)
constant	6.263	-4.428*	-0.438	-1.840*
	(14.680)	(2.080)	(1.100)	(0.750)
N.Obs	3910	3910	3910	3910
Log-Likelihood		-884.014	-53243.556	-5313.814

+ p<0.10, * p<0.05, ** p<0.01

Interview questions

- Breakthrough
  - In the period of 1997-1998 were you and your peers aware that a breakthrough was about to be discovered? Was there excitement due to a potential impactful discovery?
  - Were scientists trying to solve a specific puzzling mechanism or did they just happen to stumble on the RNAi mechanism by chance while looking for something else?
  - Were there many teams working towards solving the same problem? Was there racing?
  - Do you feel like the breakthrough could have been made earlier? What was the missing link that prevented it?
  - Was the discovery and its results a surprise? In terms of simplicity or complexity of the solution, in terms of who made the discovery?
  - Before you chose your research direction, how do you evaluate the potential impact of your research? How?
  - What papers or findings spurred your interest in RNAi research? What works had a decisive influence on your research interests?
  - What experiments, field or prior breakthroughs do you believe paved the road to the discovery? What inventions (tools), environment fostered the discovery?
  - Were you aware of the similar co-suppression and quelling results obtained in plants and fungi? / As a plant scientist did you think that co-suppression and quelling would be present in animals?

Interview questions (ctd.)

- Community
  - Was there a defined community of RNAi scientists prior to breakthrough?
  - How would you define the community of RNAi scientists prior to breakthrough? Which subfields of biology came together to form such a community?
  - How would you characterize this community? Social, open or collective?
  - How open was the community of scientists working towards solving this discovery? Was there an informal group established that frequently communicated and shared their ideas? Or were results withheld?
  - What kind of conference/research seminars did you attend at the time, was it phenomenom-based, organism-based or something else?
  - How do you think about conferences? What role do conferences play in your research?
  - In your opinion, did the breakthrough come from within the community or from outside?
  - In your opinion, who were the big contenders in the community to discover the mechanism to RNAi? Why?
- Diffusion and commercialization
  - What makes a breakthrough widely recognized and used?
  - How would you describe the speed of commercialization for the RNAi?
  - How do you decide whether or not to patent a piece of research?
- Theory Building
  - Explain the theoretical perspectives of breakthrough emergence in the social science literature, and ask whether and why they agree or disagree?
  - According to your experience on the patterns of discovery, do you think the theories I presented are comprehensive or whether alternative explanations exist?