

ÉCONOMIE DE LA SCIENCE

PHILIPPE AGHION – 22/11/16



COLLÈGE
DE FRANCE
— 1530 —

PARTIE 2 : LE GÉNOME HUMAIN

Heidi Williams (MIT)



**COLLÈGE
DE FRANCE**
— 1530 —

INTRODUCTION

- *Intellectual Property Rights and Innovation: Evidence from the Human Genome (JPE, 2013)*
- Les politiques de propriété intellectuelle visent à promouvoir la R&D
 - Axe traditionnel : rôle de la propriété intellectuelle dans l'incitation à la R&D (rentes de l'innovation)
 - Cependant, l'innovation est parfois cumulative, le développement de produits résultant de plusieurs étapes d'invention et de recherche
- Question : La propriété intellectuelle sur les découvertes existantes entrave-t-elle la R&D cumulative?
 - Preuves empiriques de la propriété intellectuelle sur le génome humain

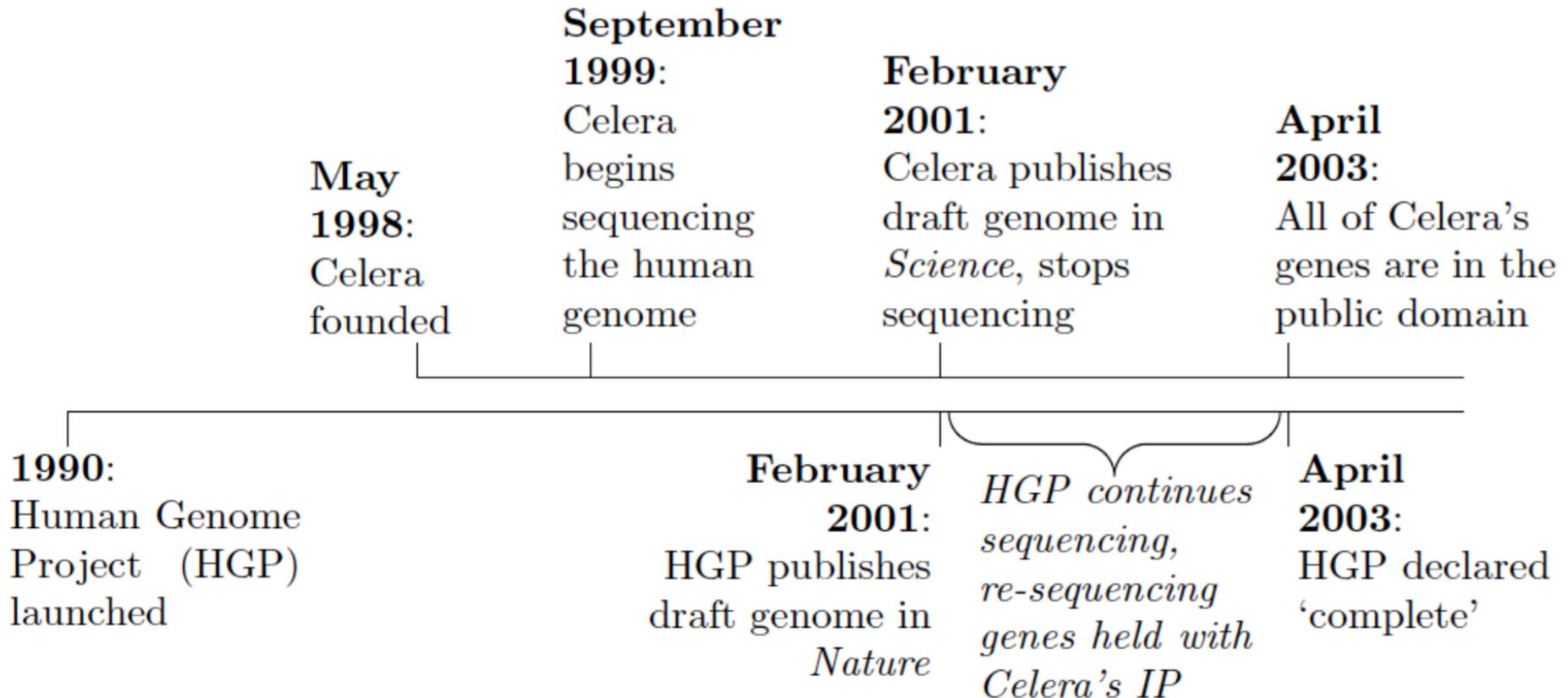


IDÉE GÉNÉRALE, UN EXEMPLE

- Supposons que la firme Celera détienne la propriété intellectuelle sur un gène humain
- Supposons que Pfizer découvre un test de dépistage d'une maladie basée sur le gène de Celera
- La propriété intellectuelle possédée par Celera découragera-t-elle Pfizer à développer ce test ?
 - Dans un environnement contractuel parfait, Celera et Pfizer négocieraient une licence de telle sorte que la recherche cumulative ne soit pas entravée
 - Cependant, les coûts de transaction (information privée sur les coûts de R&D) peuvent entraîner la rupture des négociations sur les licences
 - Riche littérature théorique sur le sujet, mais preuves empiriques rares jusqu'alors



CONTEXTE : SÉQUENÇAGE DU GÉNOME HUMAIN



QUELQUES POINTS SCIENTIFIQUES

- Gènes: portion d'ADN permettant de générer des protéines. L'ensemble de l'information génétique constitue le génotype.
 - Les protéines induisent une variation des phénotypes (caractéristiques observables)
 - Les gènes peuvent être impliqués dans de multiples liaisons génotype-phénotype
 - Un lien génotype-phénotype peut impliquer plus d'un gène
- Les gènes sont utilisés comme unité d'analyse



DÉMARCHE

- Construction d'une base de données permettant le suivi temporel (i) du séquençage des gènes, et (ii) de la propriété intellectuelle de Celera
 - Suivre l'innovation cumulée avec les liens entre phénotype et génotype
 - Pour chaque gène, collecte de données dans la base *Online Mendelian Inheritance in Man* (OMIM) sur :
 - Publications concernant l'étude de ce gène
 - Avancées scientifiques sur les liens génotype-phénotype
 - Développement de tests génétiques



EXEMPLE

- Gène RAX2
- Séquencé pour la première fois en 2001
- Celera n'en a jamais eu la propriété intellectuelle
- Publication :
 - *Human Molecular Genetics (2004)*
- Liens génotype-phénotype :
 - Dégénérescence maculaire liée à l'âge (DMLA)
 - Dystrophie de la tige du cône
- *Développement de tests génétiques :*
 - *Test de dépistage de la DMLA depuis 2009*



STATISTIQUES DESCRIPTIVES

	mean	med	std dev	min	max
Panel A: Sequencing & Celera's IP					
year sequence disclosed	2002.962	2001	3.551	1999	2009
1(Celera gene)	0.060	0	0.238	0	1

Panel B: Outcome variables					
publications in 2001-2009	2.197	0	9.133	0	231
1(known, uncertain phenotype)	0.453	0	0.498	0	1
1(known, certain phenotype)	0.081	0	0.273	0	1
1(used in any diagnostic test)	0.060	0	0.238	0	1

$N = 27,882$

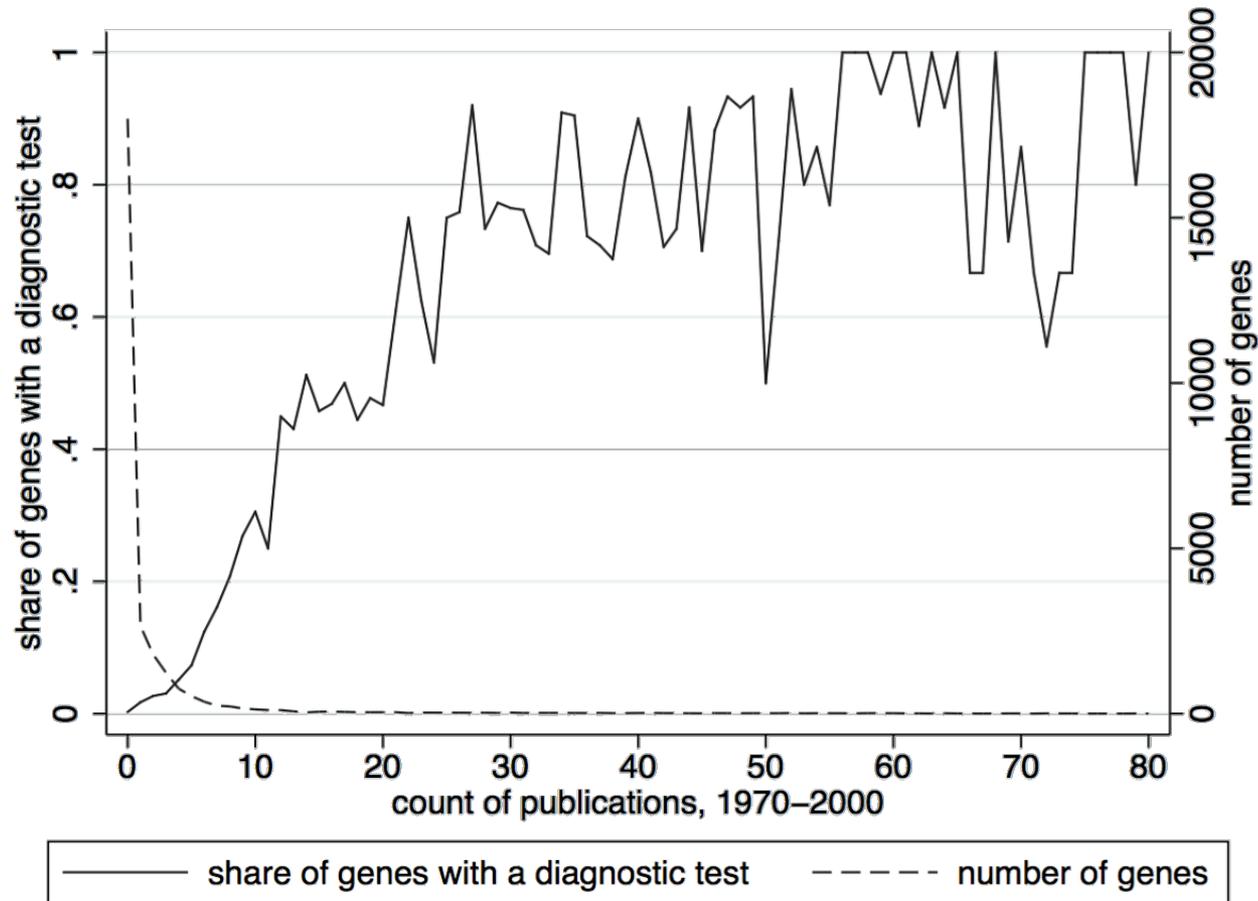


RECHERCHE SCIENTIFIQUE ET DÉVELOPPEMENT DE PRODUITS

- Littérature antérieure centrée sur les résultats liés à la publication
- Cette étude s'intéresse à la fois aux publications et aux produits commerciaux
- A priori, la distinction est importante car si les publications scientifiques ont une incitation forte à la divulgation de résultats, ce n'est pas le cas pour les produits commerciaux
- Toutefois, les deux graphiques suivants montrent que, dans le cadre de cette étude, on a un lien entre recherche scientifique et développement de produits



RECHERCHE SCIENTIFIQUE ET DÉVELOPPEMENT DE PRODUITS



La proportion de gènes utilisés dans un test augmente avec le nombre de publications



CONTEXTE EMPIRIQUE

- Human Genome Project (HGP)
 - Soumissions des nouvelles découvertes chaque jour en ligne
- Celera : Droit d'auteur / droit contractuel
 - Accès aux données gratuit pour les universitaires, Aucune restriction de publication
 - Mais restrictions à la redistribution des données pour les entreprises : discrimination par les prix (entre 5 et 15 millions de dollars par an)
- Propriété intellectuelle de Celera supprimée lorsque les gènes sont séquencés par HGP



PREMIERS RÉSULTATS

	(1) Celera mean	(2) Non-Celera mean	(3) difference [(1)-(2)]	(4) <i>p</i> -value of difference
publications in 2001-2009	1.239	2.116	-0.877	[0.000]
1(known, uncertain phenotype)	0.401	0.563	-0.162	[0.000]
1(known, certain phenotype)	0.046	0.073	-0.027	[0.000]
1(used in any diagnostic test)	0.030	0.054	-0.023	[0.000]
<i>N</i>	1,682	2,851		

La propriété intellectuelle de Celera a entraîné des réductions économiquement et statistiquement significatives de la R&D ultérieure, de l'ordre de 20 à 30%



RÉGRESSION (1) : CROSS-SECTION

- On regarde pour *l'outcome* pour chaque gène :

$$(outcome)_g = \beta(celera)_g + \lambda'(covariates)_g + \epsilon_g$$

- Observation: gène (g)
- Modèles de moindres carrés ordinaires (OLS)
- Dans les tableau de résultats :
 - Colonne 1 : régression simple
 - Colonne 2 et 3 : tests de robustesse



RÉSULTATS (1) : PUBLICATIONS

	(1)	(2)	(3)
Panel A: publications in 2001-2009			
	mean = 2.197		
<i>celera</i>	-0.877 (0.177)***	-0.328 (0.099)***	-0.264 (0.107)***
indicator variables for year of disclosure	yes	yes	yes
number of publications in each year 1970-2000	no	yes	yes
detailed cytogenetic & molecular covariates	no	no	yes
<i>N</i>	27,882	27,882	16,485

Les gènes dont la propriété intellectuelle a été possédée par Celera ont eu 0,88 publications de moins entre 2001 et 2009, soit une baisse de l'ordre de 40%



RÉSULTATS (1) : PHÉNOTYPE INCERTAIN

	(1)	(2)	(3)
Panel B: 1(known, uncertain phenotype)			
mean = 0.453			
<i>celera</i>	-0.162 (0.015)***	-0.158 (0.015)***	-0.128 (0.017)***
indicator variables for year of disclosure	yes	yes	yes
number of publications in each year 1970-2000	no	yes	yes
detailed cytogenetic & molecular covariates	no	no	yes
<i>N</i>	27,882	27,882	16,485

Les gènes dont la propriété intellectuelle a été possédée par Celera ont 16 points de pourcentage de chance de moins d'avoir un lien phénotypique connu mais incertain, soit une baisse de l'ordre de 35%



RÉSULTATS (1) : PHÉNOTYPE CERTAIN

	(1)	(2)	(3)
Panel C: 1(known, certain phenotype)			
mean = 0.081			
<i>celera</i>	-0.027 (0.007)***	-0.017 (0.006)***	-0.014 (0.007)**
indicator variables for year of disclosure	yes	yes	yes
number of publications in each year 1970-2000	no	yes	yes
detailed cytogenetic & molecular covariates	no	no	yes
<i>N</i>	27,882	27,882	16,485

Les gènes dont la propriété intellectuelle a été possédée par Celera ont 2,7 points de pourcentage de chance de moins d'avoir un lien phénotypique connu et certain, soit une baisse de l'ordre de 33%



RÉSULTATS (1) : UTILISATION DANS DES TESTS

	(1)	(2)	(3)
Panel D: 1(used in any diagnostic test)			
mean = 0.060			
<i>celera</i>	-0.023 (0.006)***	-0.014 (0.005)***	-0.013 (0.006)**
indicator variables for year of disclosure	yes	yes	yes
number of publications in each year 1970-2000	no	yes	yes
detailed cytogenetic & molecular covariates	no	no	yes
<i>N</i>	27,882	27,882	16,485

Les gènes dont la propriété intellectuelle a été possédée par Celera ont 2,3 points de pourcentage de chance d'être utilisés dans un test de dépistage, soit une baisse de l'ordre de 38%



RÉGRESSION (2) : PANEL

- On regarde *l'outcome* en suivant un gène dans le temps:

$$(outcome)_{gy} = \delta_g + \gamma_y + \beta(celera)_{gy} + \epsilon_{gy}$$

- Observation: gène*année (gy)
- Modèles de moindres carrés ordinaires (OLS)
- Cela nous permet de suivre l'évolution plus précisément dans le temps



RÉSULTATS (2) : PUBLICATIONS

	(1)	(2)	(3)
<u>publications</u> mean = 0.244			
<i>celera</i>	-0.160 (0.017)***	-0.121 (0.011)***	-0.109 (0.011)***
year fixed effects	yes	yes	yes
indicator variables for year of disclosure	yes	yes	no
number of publications in each year 1970-2000	no	yes	no
gene fixed effects	no	no	yes
<i>N</i>	250,938	250,938	250,938

Les gènes Celera ont eu 0,11 publications de moins par an, soit une baisse de l'ordre de 45%



RÉSULTATS (2) : PHÉNOTYPE INCERTAIN

	(1)	(2)	(3)
<u>1(known, uncertain phenotype)</u> mean = 0.381			
<i>celera</i>	-0.163 (0.009)***	-0.160 (0.008)***	-0.083 (0.008)***
year fixed effects	yes	yes	yes
indicator variables for year of disclosure	yes	yes	no
number of publications in each year 1970-2000	no	yes	no
gene fixed effects	no	no	yes
<i>N</i>	250,938	250,938	250,938

Baisse de l'ordre de 22%

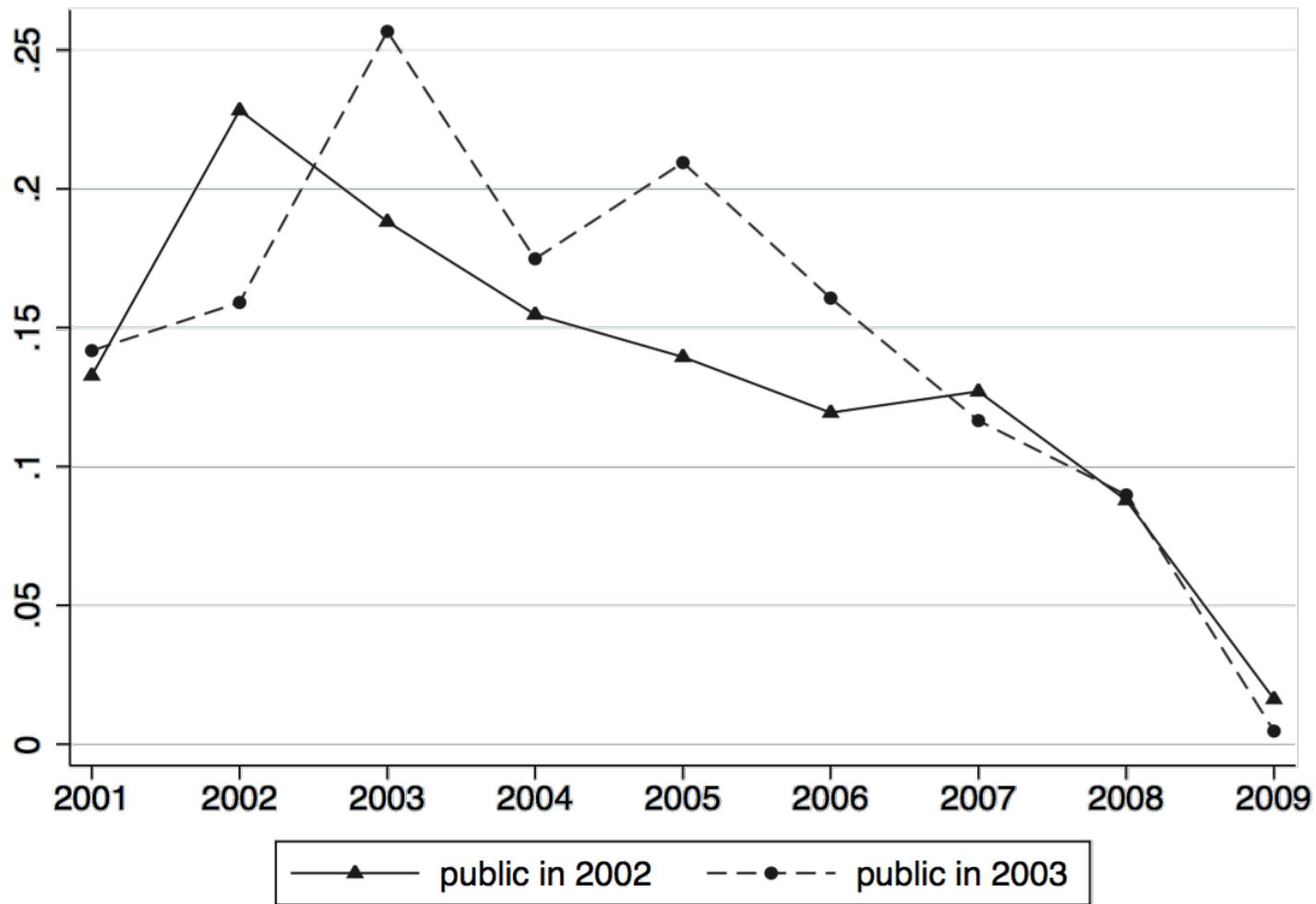


GÈNES CELERA

- On se concentre sur environ 1600 gènes dont la propriété intellectuelle a appartenu durant un certain temps à Celera
- Re-séquencé par le secteur public en 2002 ou 2003



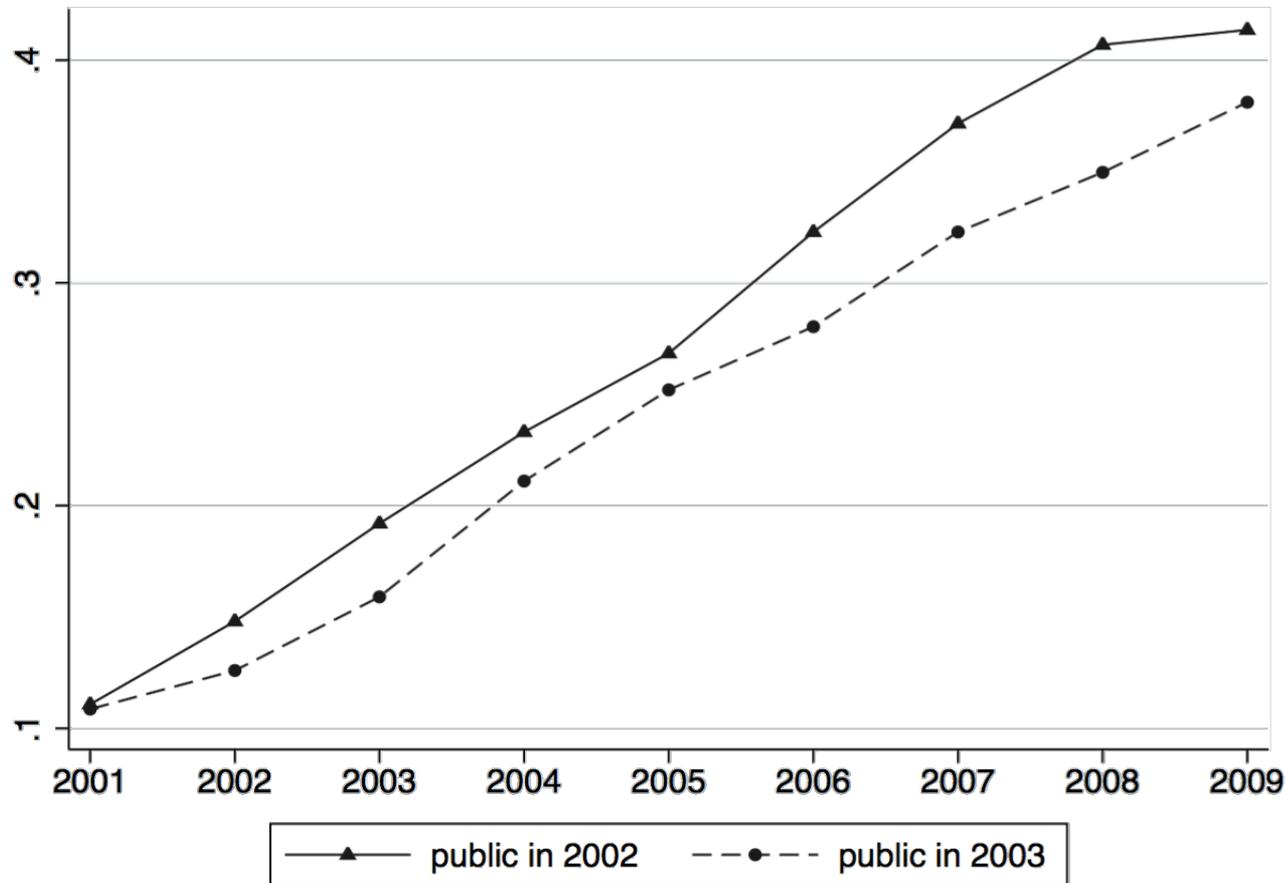
GÈNES CELERA : PUBLICATIONS



Le flux d'efforts scientifiques, mesuré par le nombre de publications, converge avec le temps entre ces deux cohortes



GÈNES CELERA : PHÉNOTYPE INCERTAIN



Le stock de connaissances scientifiques est constamment plus faible pour la cohorte de 2003



CONCLUSIONS

- On note, dans le cas du séquençage du génome humain, un **effet de la propriété intellectuelle sur la R&D**
 - Effet à court terme et à long terme
- Quel serait l'effet de la propriété intellectuelle dans d'autres secteurs que celui de la biologie ?

