Patents

https://doi.org/10.1038/s41587-026-03366-z

Scientific team structures behind FDA-approved drugs

Check for updates

A comprehensive analysis of the patents and papers behind drugs approved by the US Food and Drug Administration (FDA) reveals that successful therapeutic innovation relies on large, experienced and geographically diverse teams concentrated in a handful of elite institutions.

Drug discovery and regulatory approval represent the gold standard for successful translational biomedical research. In the United States, patents listed in the Food and Drug Administration's *Approved Drug Products with Therapeutic Equivalence Evaluations* – commonly known as the Orange Book – are those that claim an approved drug product or its use, and their inclusion requires rigorous evidence of safety and efficacy generated during clinical development[1,2].

Contemporary biomedical innovation operates at the intersection of scientific discovery and technological invention[3,4]. A considerable proportion of patents underlying FDA-approved therapeutics directly reference non-patent literature, primarily peer-reviewed scientific publications, underscoring the enduring role of upstream academic and public-sector research in producing commercially viable medicines. Because drug sponsors are legally obligated to list only patents that legitimately cover the approved product or method of use, Orange Book patents – and the scientific papers they cite, which we term upstream papers – constitute a distinctive, high-confidence subset of contributions that have demonstrably advanced an approved therapy. The emergence of comprehensive open datasets documenting patent-to-paper citations[5,6], together with disambiguated records of authors, inventors and institutions, now permits precise attribution of the research teams and individuals whose work has directly informed FDA-approved medicines.

Although team science now dominates biomedicine, with authorship lists lengthening dramatically[7,8] and inter-institutional and international collaboration becoming commonplace[7,9,10], the specific team configurations associated with research that achieves regulatory approval have remained largely uncharacterized. Women, despite comprising roughly half of biomedical PhD graduates, continue to be underrepresented in senior authorship positions on high-impact publications[11,12] and among principal investigators on industry-sponsored clinical trials[13,14]. Whether these broad trends are mirrored, exaggerated or reversed in the narrow subset of science and invention that underpins approved drugs has been unknown.

Here, we analyze 9,690 Orange Book–listed patents granted between 1985 and 2023, along with 80,176 unique upstream scientific papers directly cited in these patents. Using matched control samples from the broader patent and publication corpora, we characterize differences in impact, disruptiveness and – critically – team composition, including gender balance, experience, interdisciplinarity and geographic dispersion. Our results reveal that successful therapeutic innovation emerges from large, experienced and geographically diverse teams, yet these teams exhibit contrasting gender patterns across the translational pipeline and remain heavily concentrated in a small number of elite institutions. These findings highlight both the strengths of current ecosystems and the potential benefits of broadening participation to enhance innovation and equity in drug development.

Methods

Our analysis draws on multiple publicly available datasets linked at the level of patents and scientific publications. We identified pharmaceutical patents using the Orange Book. This resource lists drug products approved on the basis of safety and effectiveness under the Federal Food, Drug and Cosmetic Act, along with associated patent and regulatory exclusivity information. Since the 1984 Hatch-Waxman Act, applicants submitting New Drug Applications have been required to declare patents claiming the drug product or its use to the FDA, and these patents are recorded in the Orange Book[15]. We relied on a comprehensive archival compilation of Orange Book patent listings covering approvals from 1985 to 2023, encompassing 9,690 unique granted US utility patents[16].

Patent-level metadata, including filing and grant years, forward and backward citation counts, claims, technology classifications and inventor details, were obtained from the US Patent and Trademark Office (USPTO) PatentsView database. PatentsView provides disambiguated inventor identifiers along with imputed gender and location information derived from an established gender disambiguation algorithm.

To identify non-patent literature references – specifically, scientific papers cited by Orange Book patents – we drew on the "Reliance on Science" dataset, which comprehensively documents patent-to-paper citations[6,17]. We linked these citations to full bibliographic records using the OpenAlex database, resulting in 80,176 unique upstream scientific papers after deduplication and validation. Our OpenAlex extract includes detailed metadata for over 2.5 million scientific publications, such as publication year, journal or venue, field of study and citation counts. Author disambiguation and institutional affiliations are provided natively by OpenAlex, whereas author gender is inferred from first names using established machine-learning classifiers[18].

For each Orange Book-listed patent, we constructed a matched control group of up to five patents granted by the USPTO that were filed in the same year and assigned to the same primary Cooperative Patent Classification (CPC) section. Similarly, for each upstream scientific paper cited by an Orange Book patent, we selected up to five control papers published in the same year, in the same journal or conference venue, and classified in the same primary field

of study according to OpenAlex. These matched controls enable comparisons that account for observable differences in timing, technology area and publication venue.

Characterization of Orange Book patents

We begin by characterizing Orange Book–listed patents. The annual number of Orange Book patents has risen substantially over time, though a recent downturn is evident, likely attributable to lags between patent grant and drug approval (Fig. 1a). Consistent with their pharmaceutical focus, these patents are overwhelmingly classified in Cooperative Patent Classification (CPC) sections A (human necessities) and C (chemistry; metallurgy).

Given that Orange Book patents underpin FDA-approved therapeutics, they are expected to demonstrate superior technological and economic significance compared with comparable patents. We assess this using several established indicators of patent impact and value: number of backward citations (reflecting breadth of prior art relied upon), forward citation percentile rank (within the same grant year and primary CPC section), disruptiveness (consolidation–destabilization index > 0, indicating disruption rather than consolidation of existing technology[19]; Supplementary Note 6), estimated economic value (based on stock market responses to patent grants), and number of claims (serving as a proxy for scope and importance). Relative to matched controls, Orange Book patents exhibit substantially more backward citations (percentile difference Δ = 31.2%), higher forward citation impact (Δ = 52.4%), greater economic value (Δ = 55.7%) and broader claim scope (Δ = 27.2%). In contrast, they are less likely to be disruptive (Δ = −39.6%; Fig. 1b–g), aligning with the incremental nature of much pharmaceutical innovation, which typically builds on established platforms rather than introducing radical breakthroughs.

We next examine inventor team composition for the 8,057 multi-inventor Orange Book patents in our sample (Fig. 1h–n; details on variable construction are provided in Supplementary Notes 5–8). Contrary to the notion that commercially oriented pharmaceutical patents might feature less diverse or more male-dominated teams, Orange Book patents show significantly higher female inventor participation (Δ = 25.9%). Their teams are also more experienced (Δ = 10.2%) and geographically dispersed, with greater likelihood of inventors from multiple locations (Δ = 11.6%) and multiple countries (Δ = 59.4%), and larger average geospatial distances between members (Δ = 19.7%). However, these teams display lower interdisciplinary diversity (Δ = −10.3%), consistent with heightened technological specialization in drug-related invention. Similar patterns emerge when focusing on team leaders (defined as the first-listed inventor): leaders of Orange Book patents are more likely to be female (Δ = 15.5%) and more experienced (Δ = 13.2%), with no meaningful difference in individual productivity ($P = 0.6$; Fig. 1o–u).

To rigorously evaluate team composition, we estimate ordinary least squares (OLS) regressions on the full USPTO patent sample, regressing each diversity measure on an Orange Book indicator while controlling for grant year, primary CPC section, backward citation count and team size fixed effects. Results are presented both with and without inventor fixed effects. Diversity metrics (gender, experience, discipline, institution and country) are calculated using Shannon entropy, standardized as *z*-scores. We find evidence that Orange Book patents are associated with greater gender and geographic diversity ($P < 0.001$ across specifications), but lower diversity in experience and field of expertise (effects are non-significant in models without inventor fixed effects).

Characterization of Orange Book upstream papers

We next characterize the scientific papers cited by Orange Book patents relative to their matched controls. Owing to the extended timelines from scientific publication to drug approval, the annual count of upstream papers peaks in the early 2000s and declines thereafter (Fig. 2a). Although contemporary upstream papers are predominantly in the health and life sciences, earlier cohorts include a substantial share from the physical sciences.

As anticipated for publications that have demonstrably informed FDA-approved therapeutics, upstream papers exhibit markedly higher scientific impact as compared with matched controls. Relative to controls, upstream papers feature substantially more references (Δ = 30.8%) and greater forward citation impact (Δ = 29.1%). In contrast to pharmaceutical patents, upstream papers are more likely to be disruptive (Δ = 13.0%). They are also more frequently open access (Δ = 11.1%) and more likely to receive US National Institutes of Health or National Science Foundation (NIH/NSF) funding (Δ = 8.8%; Fig. 2b–g).

We then examine author team composition for the 54,010 multi-author upstream papers in our sample (Fig. 2h–n). In contrast to the pattern observed for Orange Book patents, upstream papers show significantly lower female author participation (Δ = −7.1%), suggestive of gender disparities in the scientific foundations of approved therapeutics. Their teams are nevertheless more experienced (Δ = 5.6%) and more interdisciplinary (Δ = 3.9%), and exhibit greater geographic dispersion, with higher likelihood of authors from multiple institutions (Δ = 7.4%) and multiple countries (Δ = 18.1%), and larger average geospatial distances between members (Δ = 10.5%). Similar patterns hold when examining team leaders (defined as the first- or last-listed author): leaders of upstream papers are less likely to be female (Δ = −12.8% for first authors; Δ = −6.7% for last authors) and likely to be more experienced (Δ = 10.9% for first authors; Δ = 6.2% for last authors; Fig. 2o–u).

We also estimate OLS regressions on the full OpenAlex sample, regressing each diversity measure on an upstream paper indicator while controlling for publication year, primary second-level field, journal or conference venue, reference count and team size. Results are presented both with and without author fixed effects. We find consistent evidence that upstream papers are associated with greater diversity in gender, experience, field of expertise and geography ($P < 0.001$ across specifications).

Geographic distribution of Orange Book patent teams

To examine the geographic distribution of inventor teams for Orange Book patents, we leverage PatentsView's disambiguated location and country metadata for each inventor to construct collaboration networks (Fig. 3a). The network reveals a highly concentrated core–periphery structure, with collaboration heavily centered on a small number of hubs. Strong trans-Atlantic linkages connect the United States to Western Europe, whereas connections to East Asia are limited primarily to Japan, and ties to the rest of the world remain sparse. Nearly all international collaborations involve the United States as a central node; direct partnerships bypassing this North Atlantic axis are rare. Within the United States, inventor contributions cluster in established biotechnology and pharmaceutical hubs. The leading regions are the West Coast (particularly California's Silicon Valley and Los Angeles areas) and the Northeast (Boston–Cambridge and New York metropolitan areas), with additional notable activity in Chicago.

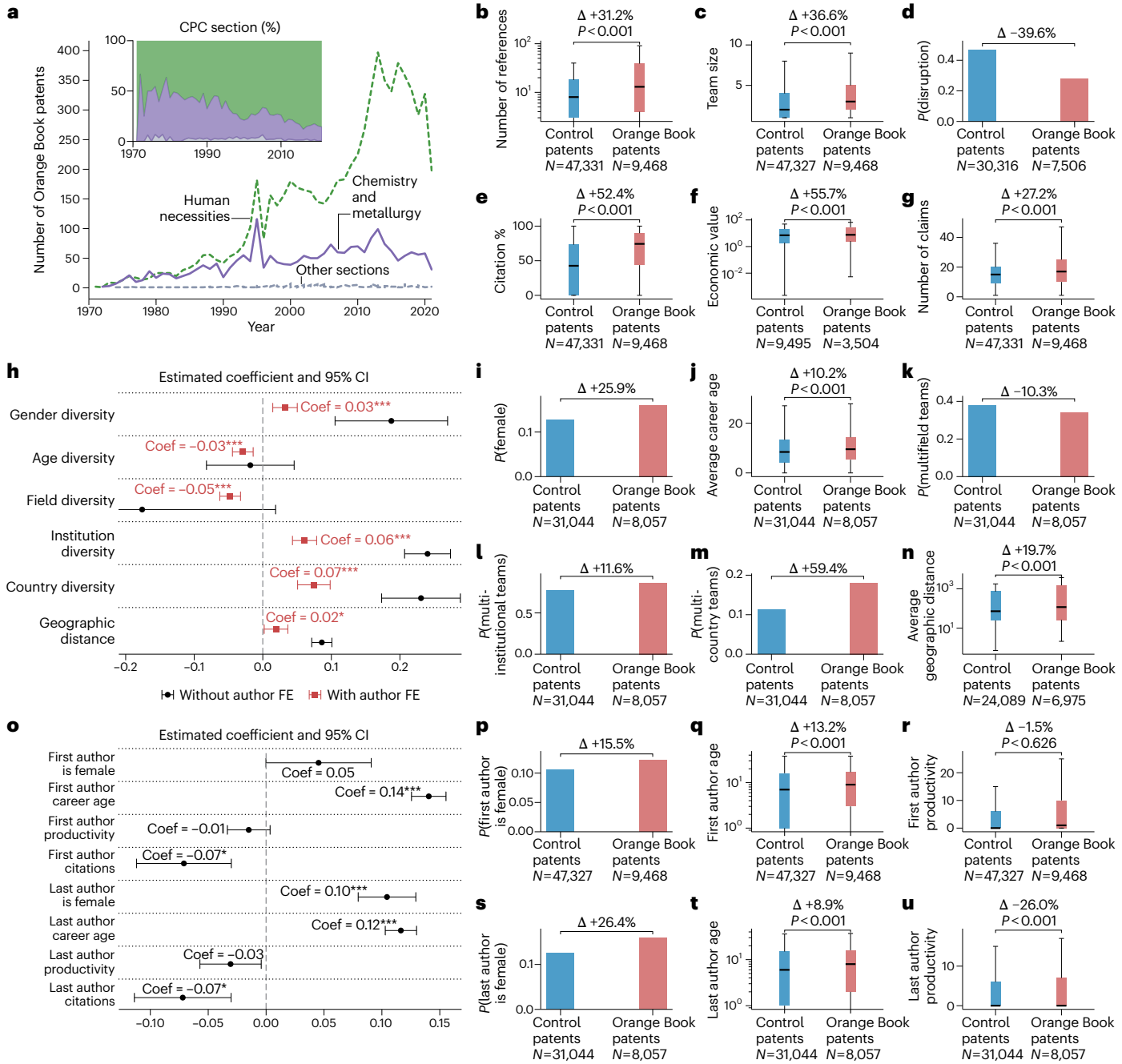


Fig. 1 | Characteristics of Orange Book patents. a, Annual trends in the number of Orange Book–listed patents by primary Cooperative Patent Classification (CPC) section. **b**–**g**, Comparisons of patent impact and value metrics between Orange Book patents and matched control patents; percentile differences (Δ) are reported, with *P* values from Welch's *t*-tests shown for non-proportion measures. **h**, Estimated coefficients (coef) from ordinary least squares (OLS) regressions with fixed effects (FE) relating an Orange Book indicator to standardized team diversity measures (Shannon entropy-based *z*-scores; see Supplementary Tables 8–12). CI, confidence interval. **i**–**n**, Comparisons of team composition characteristics between Orange Book patents and matched controls. **o**, Estimated coefficients from OLS fixed effects regressions relating an Orange Book indicator to team leadership characteristics. **p**–**u**, Comparisons of leadership characteristics between Orange Book patents and matched controls.

Of the 12,217 unique inventors associated with Orange Book patents, 6,911 (56.6%) are located in the United States, far outpacing other countries: Japan (1,055 inventors), Germany (844) and the United Kingdom (797; Fig. 3b). At the state level, California leads decisively with 1,773 inventors, followed by New Jersey (947), Massachusetts (825), Pennsylvania (554), Illinois (382) and New York (357; Fig. 3c). These concentrations align closely with the geographic footprints of major pharmaceutical firms and leading academic medical centers.

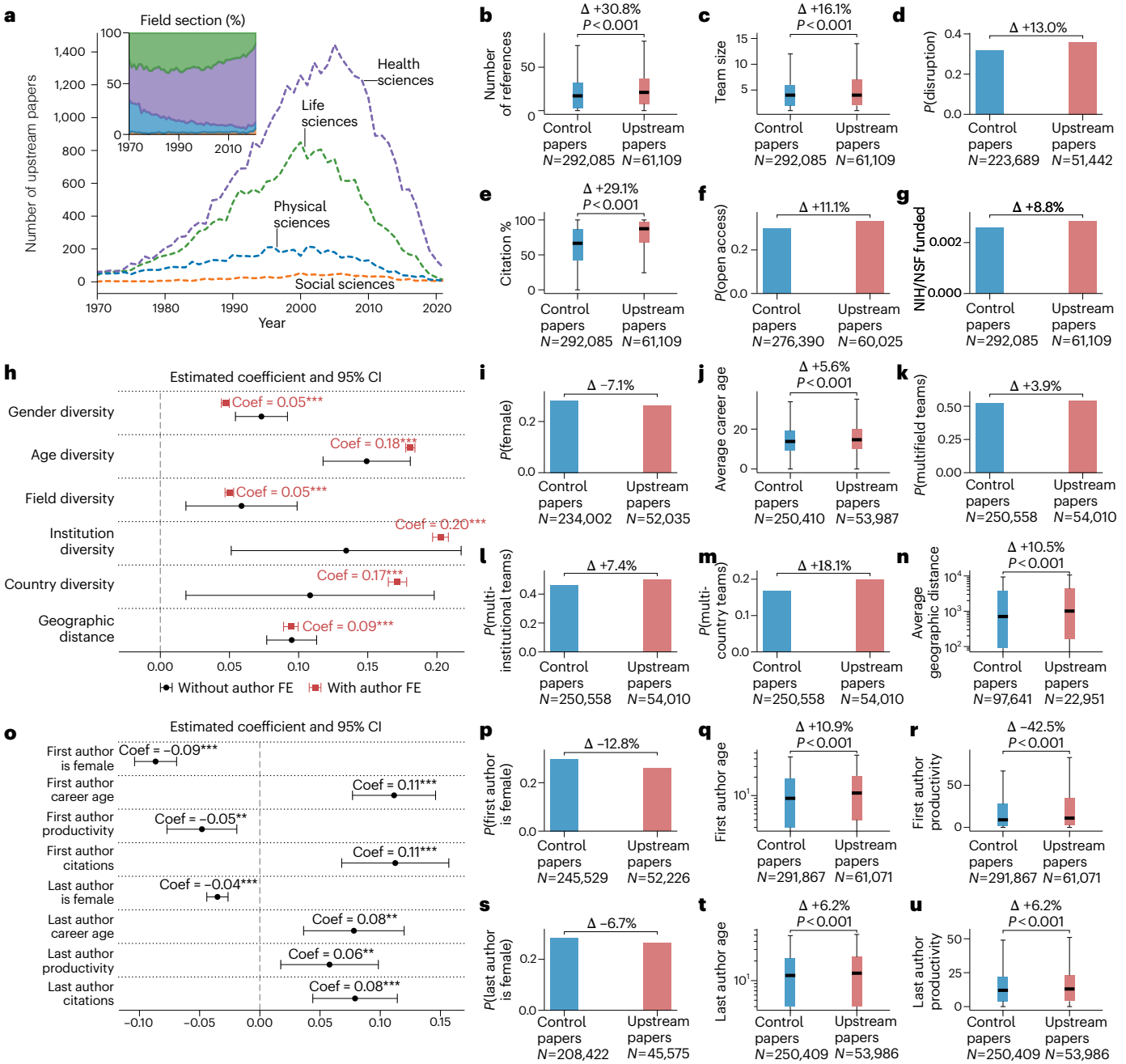


Fig. 2 | Characteristics of Orange Book upstream papers. a, Annual trends in the number of upstream papers by primary field of study. **b**–**g**, Comparisons of characteristics between upstream papers and matched control papers; percentile differences (Δ) are reported, with *P* values from Welch's *t*-tests indicated for continuous measures. **h**, Estimated coefficients from OLS regressions with fixed effects (FE) relating an upstream paper indicator to standardized team diversity measures. **i**–**n**, Comparisons of team composition characteristics between upstream papers and matched controls. **o**, Estimated coefficients (coef) from OLS regressions with fixed effects relating an upstream paper indicator to team leadership characteristics (defined as first- and last-listed authors). **p**–**u**, Comparisons of leadership characteristics (first- and last-listed authors) between upstream papers and matched controls.

We observe similar centralization at the assignee level. Among organizations holding Orange Book patents, a handful of firms dominate: Novartis (169 patents), Alza (110), Antecip Bioventures II (109), Eli Lilly (106), Pfizer (101) and Allergan (100). This oligopolistic structure highlights the critical role of large pharmaceutical companies in translating research into approved therapeutics, with many of these firms headquartered or maintaining major R&D facilities in the same high-contribution states (for example, New Jersey for Novartis and Pfizer, California for Allergan and Alza, Massachusetts for many biotech clusters).

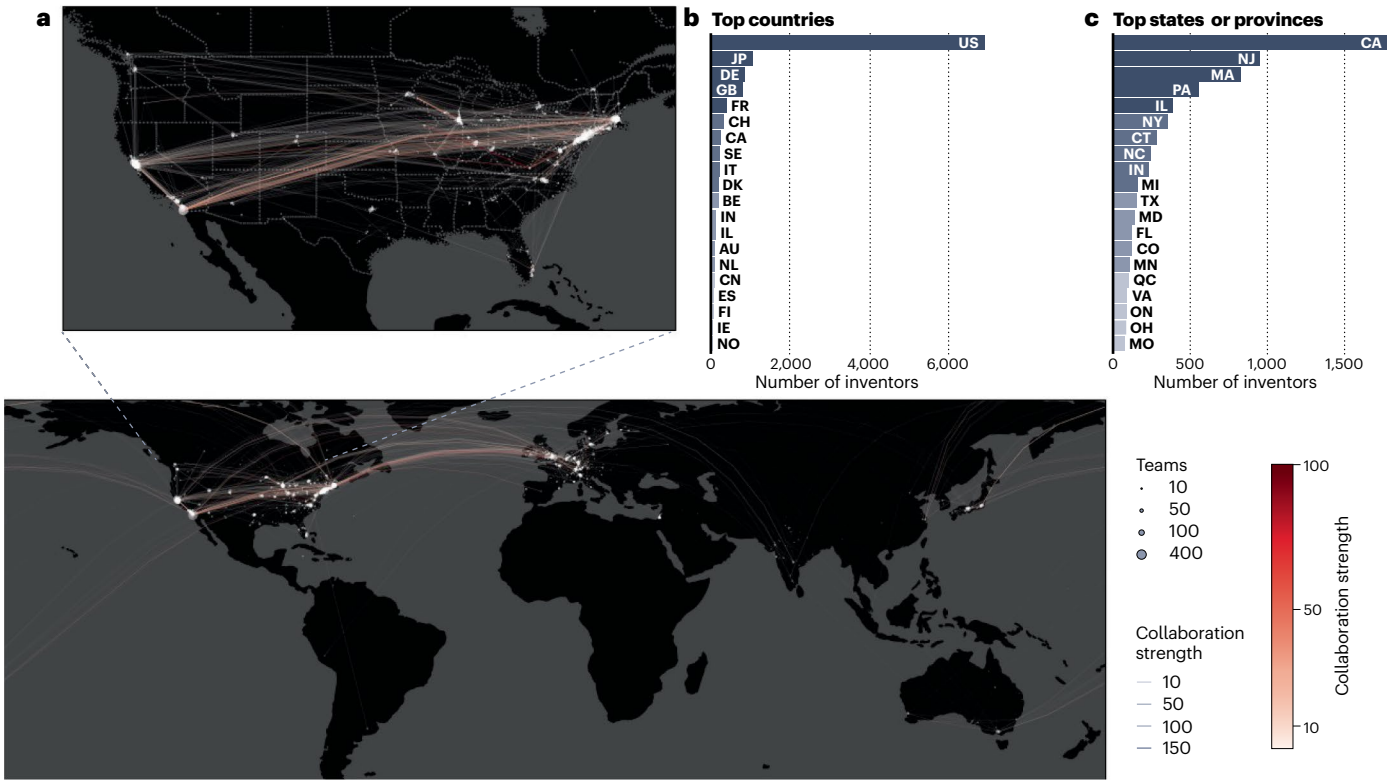


Fig. 3 | Geographic and institutional concentration of Orange Book patent teams. a, Global collaboration network overlaid on a world map with a zoom on the United States. Nodes are positioned at actual geographic coordinates and sized proportionally to the number of inventors affiliated with each location (fine-grained city level from PatentsView); edges are weighted by the number of co-inventorship ties between locations, with thickness and a white-to-red color gradient indicating collaboration intensity (thicker and darker edges denote more frequent collaborations). **b**, Ranked contributions of the top 20 countries by number of unique inventors associated with Orange Book patents. US, United States; JP, Japan; DE, Germany; GB, United Kingdom; FR, France; CH, Switzerland; CA, Canada; SE, Sweden; IT, Italy; DK, Denmark; BE, Belgium; IN, India; IL, Israel; AU, Australia; NL, Netherlands; CN, China; ES, Spain; FI, Finland; IE, Iceland; NO, Norway. **c**, Ranked contributions of the top US states by number of unique inventors associated with Orange Book patents.

These geographic patterns suggest that proximity to established industry ecosystems — encompassing corporate R&D centers, venture capital and top-tier universities — continues to confer substantial advantages in producing commercially successful pharmaceutical innovation. The pronounced US dominance and limited direct international linkages outside the North Atlantic corridor may reflect regulatory alignment (FDA-centric approvals), shared language and legal frameworks, and historical investment concentrations. The relative underrepresentation of inventors from emerging biotechnology regions points to potential untapped opportunities for broadening the geographic diversity of teams contributing to future drug development.

Geography distribution of Orange Book patents team members

To further characterize the geographic distribution of author teams for Orange Book upstream papers, we use OpenAlex's disambiguated institution, city and country metadata to construct collaboration networks and quantify contributions at multiple levels. As observed for Orange Book patents, the network exhibits a pronounced core-periphery structure dominated by US institutions (Fig. 4a). Strong trans-Atlantic connections link the United States to Western Europe, whereas ties to East Asia and the rest of the world remain limited. Nearly all international collaborations involve a US institution as an intermediary; direct partnerships outside the North Atlantic corridor are exceedingly rare. Within the United States, authorship concentrates in the Northeast (Boston–Cambridge area) and West Coast (California), home to many leading research universities and medical centers.

Of the 185,407 unique authors contributing to upstream papers, 90,767 (49.0%) are affiliated with US institutions, far exceeding other countries: the United Kingdom (16,101), Japan (13,194) and Germany (11,306). Despite China's rapid growth in overall scientific output, it ranks only eighth (4,450 authors), behind Canada (7,160). The top 10 countries account for approximately 90% of all authorships and the top 20 for 99%, illustrating the extreme geographic concentration in the scientific foundations of FDA-approved therapeutics (Fig. 4b).

Institutional contributions are even more skewed (Fig. 4c). Harvard University leads with 3,893 affiliated authors, followed by the National Institutes of Health (2,865), the French National Institute of Health and Medical Research (INSERM; 2,322), Johns Hopkins University (2,213) and the University of California, San Francisco (2,003). The top 10 institutions alone represent roughly 12% of the entire author corpus.

This marked geographic and institutional concentration highlights the outsized role

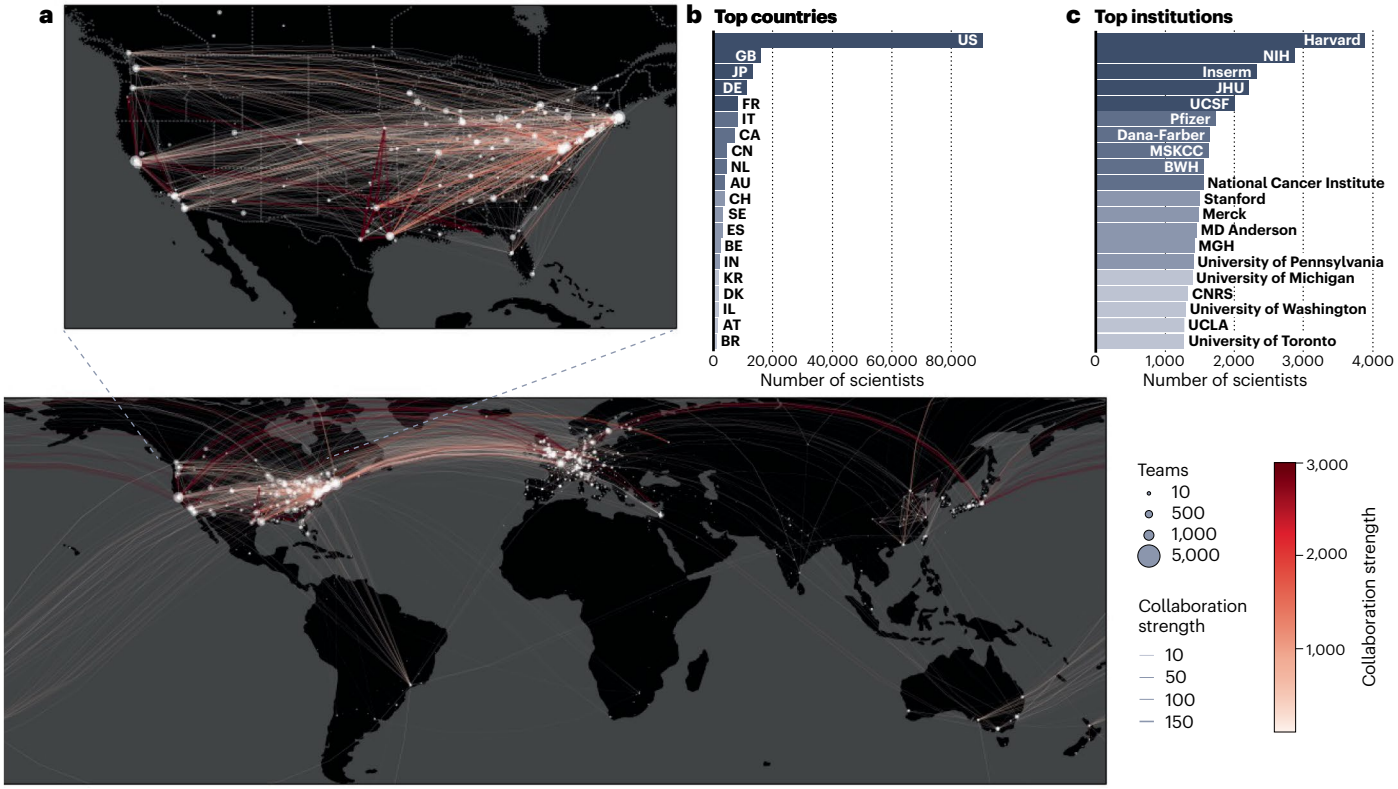


Fig. 4 | Geographic and institutional concentration of Orange Book upstream paper teams. a, Global collaboration network overlaid on a world map with a zoom on the United States. Nodes are positioned at actual geographic coordinates and sized proportionally to the number of authors affiliated with each institution (from OpenAlex); edges are weighted by the number of co-authorship ties between institutions, with thickness and a white-to-red color gradient indicating collaboration intensity (thicker and darker edges denote more frequent collaborations). **b**, Ranked contributions of the top 20 countries by number of unique authors associated with upstream papers. KR, South Korea; AT, Austria; BR, Brazil; others as in Fig. 3b. **c**, Ranked contributions of the top 20 institutions by number of unique affiliated authors contributing to upstream papers. JHU, Johns Hopkins University; UCSF, University of California, San Francisco; MSKCC, Memorial Sloan Kettering Cancer Center; BWH, Brigham and Women's Hospital; MGH, Massachusetts General Hospital; CNRS, Centre National de la Recherche Scientifique; UCLA, University of California, Los Angeles.

of a small number of elite US academic and government research entities — supported by substantial public funding and embedded in dense knowledge ecosystems — in generating the upstream science that ultimately informs approved drugs. The persistence of strong trans-Atlantic ties reflects historical collaborations, shared linguistic and funding networks, and alignment with US regulatory priorities. The relative underrepresentation of authors from high-output nations such as China may stem from differences in research focus (for example, less emphasis on translational biomedical areas relevant to FDA approvals), language barriers in publication, or limited integration into global clinical and pharmaceutical networks. These patterns suggest that, despite the globalization of science, the foundational research underpinning US-approved therapeutics remains heavily reliant on a narrow geographic and institutional base, potentially limiting the diversity of perspectives and approaches that feed into drug development pipelines. Broadening international and institutional participation in upstream research could enhance the robustness and inclusivity of translational outcomes.

Discussion

Our analysis of Orange Book–listed utility patents and upstream scientific papers that directly inform FDA-approved therapeutics offers a detailed empirical portrait of the team structures and geographic patterns associated with successful translational biomedical research[20,21]. The results paint a consistent picture: translational impact is linked to large, experienced and geographically dispersed teams, yet these teams are overwhelmingly anchored in a narrow set of elite institutions and locations.

These findings both corroborate and extend previous bibliometric analyses of high-impact biomedical research[8,22]. The positive association with larger team size reflects the broader rise of 'big team science' in biomedicine, where collaborative scale facilitates the resource-intensive, multiphase work required for drug development. The greater institutional and geographic diversity likely captures the benefits of pooling specialized expertise, shared infrastructure and regulatory familiarity across complementary centers.

A striking divergence emerges along the translational pipeline with respect to innovation type. Orange Book patents tend to be consolidating rather than disruptive, building incrementally on existing technological platforms — a pattern consistent with the applied, development-oriented goals of pharmaceutical invention. Upstream papers, however, are significantly more disruptive than matched

controls (Welch's t-test, $P < 0.001$). This contrast highlights a classic division of labor in translational science: foundational discoveries often arise from disruptive academic insights that reframe problems or introduce novel mechanisms, whereas successful commercialization requires subsequent consolidation, refinement and scaling within more constrained technological and regulatory boundaries. Our results thus underscore the complementary roles of disruptive 'science push' and consolidating 'demand pull' in driving therapeutic advances[3].

Gender composition reveals another complex pattern. Orange Book patents exhibit substantially higher female inventor participation than comparable patents, with women more prominently represented both on teams and in leadership positions (as first-listed inventors). This suggests that, at the stage of applied invention and drug development, female innovators play a disproportionately important role in producing commercially successful outcomes. In contrast, upstream scientific papers show significantly lower female authorship (Welch's t-test, $P < 0.001$), particularly in leadership roles (first and last authors). This disparity indicates that gender imbalances persist in the foundational academic research that feeds the pipeline, even as women appear to gain greater representation downstream during translational and commercialization phases. The shift may reflect differences in career incentives, risk tolerance or selection into industry versus academia, but it also raises the possibility that translational success actively benefits from greater gender diversity at later stages. Closing persistent gaps in senior academic authorship remains essential to ensure the full talent pool contributes to upstream discoveries.

The extreme geographic and institutional concentration documented here carries important implications. Although multi-institutional and international collaboration is overrepresented among translationally successful work, such diversity is almost exclusively confined to an established network of flagship entities. This 'elite club' structure suggests that historical accumulations of institutional capital — dense ecosystems of funding, talent, clinical infrastructure and regulatory expertise — continue to confer decisive advantages. Emerging biotechnology hubs in Asia, Latin America and elsewhere remain markedly underrepresented, despite rapid growth in overall scientific output. This skew not only limits the demographic and intellectual diversity of perspectives informing new therapeutics but may also constrain the pipeline's ability to address global health needs, such as diseases prevalent in underrepresented regions. Moreover, heavy reliance on a narrow geographic base introduces systemic vulnerabilities, including sensitivity to regional funding fluctuations or disruptions in international collaboration.

Our study has several limitations. First, our upstream corpus is limited to papers explicitly cited in Orange Book patents and thus captures only a fraction of influential science; earlier foundational work or contributions never formally referenced in patent documents are inevitably underrepresented. Second, gender imputation from names, while using state-of-the-art classifiers, introduces measurement error, especially for authors from underrepresented regions. Third, proxies for experience and career age rely on indexed USPTO and OpenAlex records and may undercount contributions in clinical-trial publications or pre-digital literature. Fourth, although our matched-control design accounts for observable confounders, unmeasured factors — such as funding scale, clinical-trial access or strategic citation practices — could still influence results. Finally, the analysis reflects historical patterns through 2023; substantial lags between publication, patenting and approval mean that emerging global shifts in research capacity and collaboration may not yet be visible in approved therapeutics.

Looking forward, our findings point to several policy and practice recommendations. First, funding agencies aiming to maximize translational yield could prioritize large consortia that combine scale, disciplinary focus and cross-institutional partnerships while actively encouraging inclusion of institutions beyond incumbent elites through targeted incentives or mandatory broader participation requirements. Second, concerted efforts to advance gender equity in academic leadership and principal investigator roles are needed to address upstream imbalances and ensure diverse talent flows through the entire pipeline. Third, initiatives to widen meaningful global participation — such as regulatory harmonization, shared clinical-trial networks, technology-transfer partnerships and long-term capacity building in emerging regions — could gradually reduce geographic skew without sacrificing near-term success rates. Such broadening would enhance resilience, equity and the pipeline's responsiveness to diverse health challenges.

Ultimately, the teams and institutions delivering FDA-approved medicines today remain highly unrepresentative of the global biomedical research enterprise. Whether this concentrated ecosystem can — or should — be sustainably widened without compromising translational efficiency represents a defining challenge for science policy in the decades ahead. Addressing it will require balancing proven pathways to success with deliberate efforts to democratize access to the networks that currently gatekeep therapeutic innovation.

Data availability

The Orange Book raw data are publicly available at https://www.fda.gov/drugs/drug-approvals-and-databases. The patent information associated with Orange Book can be downloaded at https://doi.org/10.7910/DVN/VI93T9. The PatentsView dataset is publicly available at https://patentsview.org/download/data-download-tables.The OpenAlex dataset is publicly available at https://docs.openalex.org/download-all-data/openalex-snapshot. The Reliance on Science dataset can be downloaded at https://relianceonscience.org/.

Code availability

The code and source data are publicly available at https://github.com/AlexJieYang/Team_FDA_Drugs.

Alex J. Yang [1] ✉, **Richard B. Freeman**[2,3] **& Sanhong Deng**[1] ✉

[1]School of Information Management, Nanjing University, Nanjing, China. [2]Center for Labor and a Just Economy, Harvard Law School, Cambridge, MA, USA. [3]National Bureau of Economic Research, Cambridge, MA, USA.
✉e-mail: alexjieyang@outlook.com; sanhong@nju.edu.cn

Published online: 10 November 2026

References

1. Durvasula, M., Hemphill, C. S., Ouellette, L. L., Sampat, B. & Williams, H. L. *Res. Policy* **52**, 104791 (2023).
2. Azoulay, P., Graff Zivin, J. S., Li, D. & Sampat, B. N. *Rev. Econ. Stud.* **86**, 117–152 (2019).
3. Ahmadpoor, M. & Jones, B. F. *Science* **357**, 583–587 (2017).
4. Fleming, L., Greene, H., Li, G., Marx, M. & Yao, D. *Science* **364**, 1139–1141 (2019).
5. Manjunath, A. et al. *Nat. Biotechnol.* **39**, 678–683 (2021).
6. Marx, M. & Fuegi, A. *J. Econ. Manag. Strategy* **31**, 369–392 (2022).
7. Jones, B. F., Wuchty, S. & Uzzi, B. *Science* **322**, 1259–1262 (2008).
8. Andalón, M., de Fontenay, C., Ginther, D. K. & Lim, K. *Nat. Biotechnol.* **42**, 1314–1319 (2024).
9. Lin, Y., Frey, C. B. & Wu, L. *Nature* **623**, 987–991 (2023).
10. Wagner, C. S., Whetsell, T. A. & Mukherjee, S. *Res. Policy* **48**, 1260–1270 (2019).

11. Koning, R., Samila, S. & Ferguson, J.-P. *Science* **372**, 1345–1348 (2021).
12. Xie, L., Zhou, J., Zong, Q. & Lu, Q. *Res. Policy* **49**, 103885 (2020).
13. Liu, F., Holme, P., Chiesa, M., AlShebli, B. & Rahwan, T. *Nat. Hum. Behav.* **7**, 353–364 (2023).
14. Ross, M. B. et al. *Nature* **608**, 135–145 (2022).
15. Ouellette, L. L. & Sampat, B. N. *JAMA Health Forum* **5**, e243775–e243775 (2024).
16. Ouellette, L. L. & Sampat, B. N. Orange Book patent data, including public-sector patent information, 1985–2023. *Harvard Dataverse* https://doi.org/10.7910/DVN/VI93T9 (2024).
17. Marx, M. & Fuegi, A. *Strateg. Manag. J.* **41**, 1572–1594 (2020).
18. Van Buskirk, I., Clauset, A. & Larremore, D. B. *Proc. Int. AAAI Conf. Weblogs Soc. Media* **17**, 866–877 (2023).
19. Funk, R. J. & Owen-Smith, J. *Manag. Sci.* **63**, 791–817 https://doi.org/10.1287/mnsc.2015.2366 (2017).
20. Chien, C. V., Halkowski, N. & Kuhn, J. *Nat. Biotechnol.* **41**, 317–321 (2023).
21. Krieger, J., Li, D. & Papanikolaou, D. *Rev. Financ. Stud.* **35**, 636–679 (2022).
22. Yang, Y., Tian, T. Y., Woodruff, T. K., Jones, B. F. & Uzzi, B. *Proc. Natl Acad. Sci. USA* **119**, e2200841119 (2022).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41587-026-03366-z.