ISSN 1831-9424

European Commission

Exploring the global landscape of biotech innovation: preliminary insights from patent analysis

Grassano, N., Napolitano, L., M'barek, R., Rodriguez Cerezo, E., Lasarte Lopez, J.

2024

Joint Research Centre

EUR 31889 EN

Contact information
Name: Robert M'barek
Address: C/ Inca Garcilaso 3; E-41092 Seville/Spain
Email: Robert.M'barek@ec.europa.eu

EU Science Hub
https://joint-research-centre.ec.europa.eu

JRC137266

EUR 31889 EN

PDF ISBN 978-92-68-13890-8 ISSN 1831-9424 doi:10.2760/567451 KJ-NA-31-889-EN-N

Luxembourg: Publications Office of the European Union, 2024

How to cite this report: European Commission, Joint Research Centre, Grassano, N., Napolitano, L., M'barek, R., Rodriguez Cerezo, E. and Lasarte Lopez, J., *Exploring the global landscape of biotech innovation: preliminary insights from patent analysis*, Publications Office of the European Union, Luxembourg, 2024, https://data.europa.eu/doi/10.2760/567451, JRC137266.

Contents

Abstract

This technical report focuses on innovation in biotechnologies (biotech), as captured by patented inventions worldwide. The study analyses international patents filed at multiple offices, at least two of which belonging to the IP5 consortium. Moreover, it relies on expert knowledge collected by the OECD to select the inventions connected to biotechnologies. With the help of experts, the different biotechnology codes have been attributed to agricultural (green biotech), industrial (white biotech) human/medical (red biotech), also adding a residual category collecting biotechnologies that have potentially a transversal use (horizontal biotech). The analysis produces a bird's-eye view on the evolution of patenting in this technological area over time and its relevance across geographical and technological dimensions. Biotechnology patents represent about 5% of total patents filed between 2001 and 2020. The vast majority of the biotechnology patents is related to industrial and medical applications, accounting for more than 96% of all patents analysed. The US is leading in the development of biotechnology patents (39% of total biotechnology patents in 2020), followed by the EU with 18% and China, which is advancing quickly (10%). Preliminary analysis suggests that the competition among countries/regions in biotechnology patents revolves around the number of patents in each of the main biotechnological domains, rather than the different kinds of biotechnologies patented.

Acknowledgements

The authors would like to thank Alexander Tübke (JRC.B6) for his support, Wim Broothaerts and Ursula Vincent (JRC.F5) for their comments on the IPC-biotechnology colour classification correspondence, Koen Jonkers (JRC.S) and two anonymous referees from JERB for their feedback on the manuscript.

Authors

Nicola Grassano

Lorenzo Napolitano

Robert M'barek

Emilio Rodríguez Cerezo

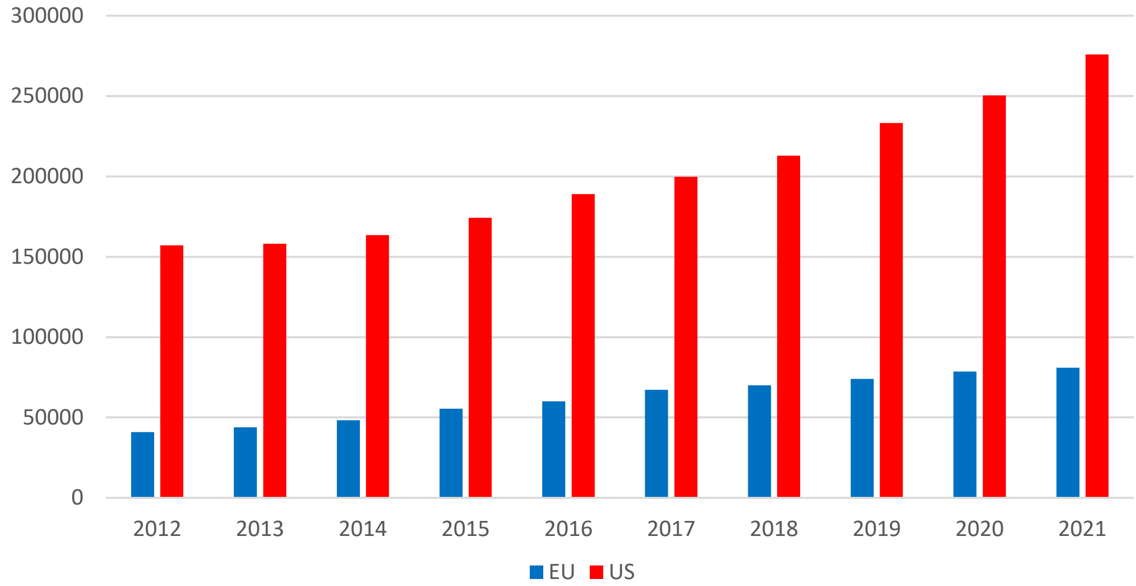
Jesús Lasarte López

1 Introduction

Research in biotechnologies (*biotech*) is a field on the rise. In the EU, the number of researchers working in biotechnology almost doubled in the last decade (from 42 000 in 2012 to 81 000 in 2021[1], approximately). Also in the United States of America (US), the absolute number of employees in biotechnology grew significantly over the same period from about 135 000 in 2012 to 250 000 in 2021, getting very close to 300 000 in 2023[2]. Thus, the scientific research sector in biotechnology employs about three times more workers in the US than in the EU (see Figure 1).

The growth rate of employment in biotechnology in the EU ranged between yearly 3% and 15% in the decade 2012-2021. For the US, the yearly growth rate of employment in biotechnology was between 1% and 11% in the period considered. The EU experienced higher yearly growth rates than the US until 2017 (except for 2016), while the US registered higher growth rates from 2018 onward.

Figure 1. Number of employees in research and development in biotechnology



Note: Data for the EU refers to the Scientific Research and Experimental development on biotechnology sector (NACE code: 7211). Data for the US refers to NAICS 541714 “Research and Development in Biotechnology (except Nanotechnology)”.
Source: Structural Business Statistics and own calculations following Ronzon et al. (2022) and US Bureau of Labor statistics[3].

Also the share of biotechnology jobs over the total of jobs in R&D has grown significantly. At the EU level, the number of jobs in biotechnology fields accounted for 17.2% of employment in the Scientific Research and Development sectors (NACE code M721) in 2021, an increase of 6 percentage points over the same value for 2012.

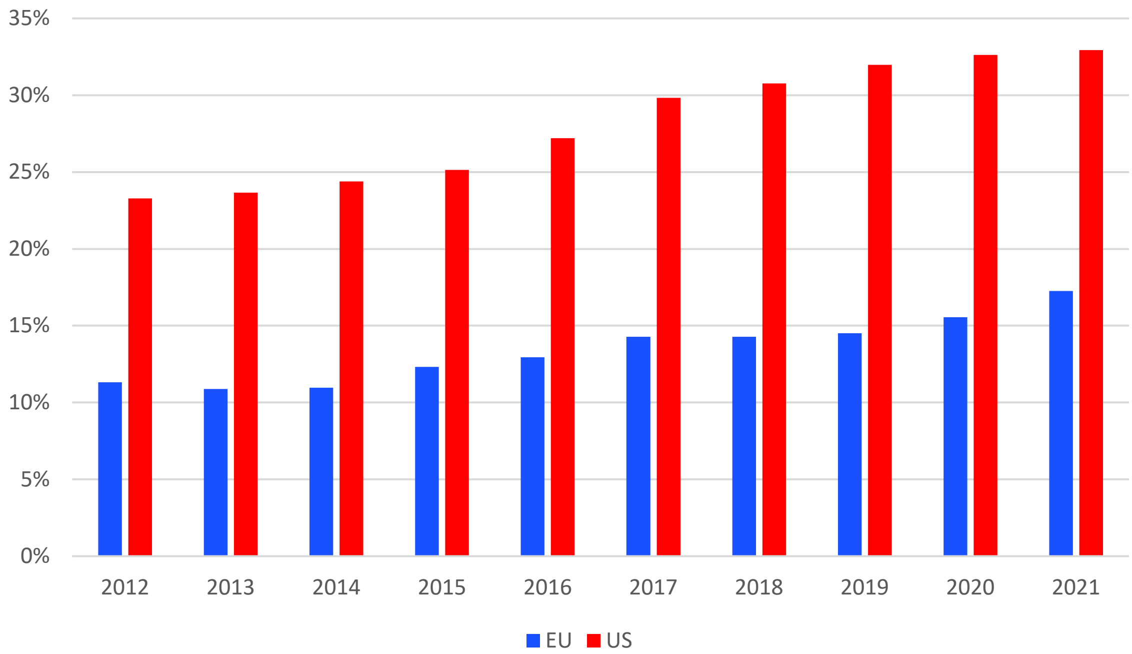
In the same time period, in the US, the importance of biotechnology employment over total employment in research and development in the physical, engineering, and life sciences has grown substantially, increasing by about 9.5 percentage points (see Figure 2).

[1] Information from Structural Business Statistics (NACE code M7211: Scientific Research and Experimental Developments on biotechnology) from Eurostat and own calculations following Ronzon *et al.* (2022).

[2] 293 216 employees in biotechnology in the US in 2023 (not shown in the chart).

[3] https://www.bls.gov/bls/about-bls.htm

Figure 2. Share of biotechnology employees over the total employment in Scientific Research and Developments on Natural Sciences and Engineering



Note: For the EU the share of biotechnology fields (NACE code M7211) over the total employment in Scientific Research and Development in Natural Sciences and Engineering (NACE 721) is considered. For the US the share of Research and Development in Biotechnology (except Nanobiotechnology) (NAICS code 541714) over the total employment in Research and Development in the physical, engineering, and life sciences (NAICS 54171) is considered.
Source: Structural Business Statistics and own calculations following Ronzon et al. (2022) and US Bureau of Labor statistics.

Against this background of rising “inputs” in the research process[4], this technical report focuses on the output of this process by analysing innovation in biotechnologies as captured by patented inventions worldwide[5]. To this aim, the report analyses international patents filed at multiple offices, at least two of which belonging to the IP5 consortium (EPO, USPTO, JPO, KIPO, CNIPA). Moreover, it relies on expert knowledge collected by the OECD to select the inventions connected to biotechnologies. The analysis aims to produce a bird’s-eye view on the evolution of patenting in this technological area over time and its relevance across the geographical and technological dimensions.

The remainder of this document is structured as follows: Section 2 describes in detail the methodology followed to identify and classify biotechnology patents according to their main use, mainly following the colour schema proposed by the European Patent Office. Section 3 presents the main trends in biotechnology patents from 2001 to 2020. Section 4 explores the comparative advantage of regions and member states in specific categories of biotechnology. Section 5 delves deeper into the analysis of biotechnology patents, providing a more detailed technological classification. Section 6 concludes.

[4] This is also true when considering private R&D investment by biotechnology firms, see Grassano *et al.* (2022).

[5] The use of patents to measure innovative output is a standard practice in innovation literature since decades (i.e. Mueller, 1966). Also the limitations of patent data have been well-known for a long time (i.e. Pavitt, 1985).

2 Building the dataset

The identification of biotechnology patents used for this analysis relies on the European Patent Office's EPO Patstat database[6], which collects comprehensive information about the patents filed over 70 national and regional patent offices, including those that process the majority of applications worldwide. PATSTAT is biannually updated by the EPO since 2007 and records information on more than 100 million patent documents grouped in over 50 million families. The database records comprehensive information about the documents including, among other things, the receiving office, the filing date, the technologies in which the patent innovates (encoded in standard technology codes), and, whenever available, the residence of the applicants and of the inventors at the time of filing.[7] Concerning temporal coverage, it is worth noting that the nature of the data collection process generates an intrinsic lag of at least 12-18 months between when an application reaches a patent office and when the corresponding record appears in Patstat. This lag increases further due to contingencies specific to individual offices[8]. Therefore, patent counts extracted from an edition of Patstat published in a year *y* are not fully reliable beyond (*y-3*) or (*y-4*).

It is worth noting that patent applicants can seek protection for an invention by submitting separate applications to a plurality of patent offices over a predetermined period. Later patent applications for the same invention can claim one or more previous ones as *priorities*, which allows establishing a chain of connected documents relating to the same invention. These allow defining so-called *families*, which can serve as the basic units of observation in empirical analyses of patenting dynamics. Usually, a family is assigned to the filing date of the first application because it is the closest recorded date to the time of conception of the invention. Patent families can be defined in more than one way. For example, a strict family definition (e.g. DOCDB patent families) may require that documents in the same families name the same set of priorities, while a looser definition (e.g. INPADOC families) may allow grouping together applications that have in common directly or indirectly at least one priority. Moreover, family definitions can be refined to select only inventions that are more likely to have high value. A common criterion to this end is to consider families for which the applicants have sought international IP protection in the more relevant patent offices worldwide. Following a consolidated approach (see for example Amoroso *et al*, 2021), this analysis focuses on INPADOC families containing applications filed at multiple patent offices, at least two of which belong to the IP5 consortium (EPO, USPTO, JPO, KIPO, CNIPA). Throughout this section, patents or patented inventions refer to IP5 families.

The identification of patents in biotechnologies exploits the fact that Patstat encodes the technological fields in which a patent innovates following two standard classifications: the International Patent Classification (IPC) and the Cooperative Patent Classification (CPC). Both classifications have a similar hierarchical structure in which 1-digit codes (i.e. *sections*) are the root of very broad areas of technology (e.g. G: *physics*); these aggregated codes branch out into progressively longer codes associated with narrower definitions (e.g. G02: *optics*; G02B: *passive optical elements*; G02B 1/02: *optical elements [...] made of crystals, e.g. rock-salt*). All seven sections included in the IPC classification feature in the CPC classification, which contains an additional section (Y) dedicated mostly to Climate-Change Mitigation Technologies[9] (CCMTs). The trees of the IPC and CPC classifications are very similar close to the roots and diverge approaching the leaves. For example, at 4-digits both classifications have around 600 unique codes; at 8-digits there are around 7000 codes in the IPC and around 10000 the CPC.

However, not all technological domains of interest (e.g. digital technologies, advanced materials, biotechnologies) have a dedicated subtree within either classification. Usually, to extract the codes that characterize a policy-relevant set of technologies from the patent classification, one must identify the relevant codes placed across the entire classification tree. In general, this non-trivial operation requires expert

[6] PATSTAT Global - 2023 Autumn Edition. Detailed information about the latest editions of the database is available at https://www.epo.org/en/searching-for-patents/business/patstat.

[7] The geographical information recorded in Patstat relates to patent applicants and inventors located in over than 200 countries. The geographical information is often incomplete, with the coverage varying widely across patent offices. The most commonly available information in this sense is the country of residence of inventors and applicants, though for some documents, data that is more granular is also available (e.g., the full address).

[8] For instance, until the early 2000s internal regulations of the US patent office (USPTO) only allowed the publication of granted patents only. This, alongside a consistent backlog of applications awaiting examination, resulted in an incomplete coverage for of up to 5 years prior to the publication year of Patstat. The USPTO has since changed its rules allowing it to compute a complete and relatively timely patent count for disclosed patent applications. However, some lag in coverage is inevitable.

[9] Specifically, the part of the CPC dedicated to tagging climate change mitigation and adaptation consists of the codes under the so-called 'Y02/Y04S' tagging scheme nested under section Y.

knowledge to produce an accurate result. For this reason, to filter patents in biotechnology, the working paper “Revised proposal for the revision of the statistical definitions of biotechnology and nanotechnology” is used, published online by the OECD. The document reports a list of 38 broader technological areas connected to biotechnology patents, which can be disaggregated into a comprehensive list of around 1000 IPC technology codes (see for details Friedrichs and van Beuzekom, 2018). To each of these 38 technological codes identified by the OECD, a “colour” according to their main field of use is attributed. Following the schema proposed by the European Patent Office[10], the 38 are divided into “red” (biotechnologies used mainly in the human/medical sector), "white" (biotechnologies primarily used in industrial processes), and "green" (biotechnologies employed in agriculture). A residual category of “horizontal” biotechnologies classifies those technologies that may be applied in different domains and is difficult to pin down to a single predominant use.

There are at least two caveats to keep in mind when reading our analysis. First, our approach is to identifying biotechnology patents through static lists of technology codes, which makes it prone to obsolescence. The technology classifications are very dynamic to accommodate the fast pace of innovation, so codes (especially the more granular ones) change over time and the same inventions can end up reclassified using new codes within a relatively short time. Furthermore, as knowledge advances new codes that were not relevant at the time of compiling the list might become important, implying that if a list of technology codes is old enough, it will account only partially for the portion of technological landscape it was devised to capture.

Second, the colour code classification of the 38 technological codes are extended to all the around 1000 IPC technology codes in which they can be disaggregated. It can well be that under a technological code classified as red there is a bus technological code that is horizontal. Moreover, to the best of the authors’ knowledge, this is the first time, the colour code classification is applied to the list of IPC technological codes identified as biotechnologies by the OECD (reported in annex 1); so the results based on this need to be treated with caution.

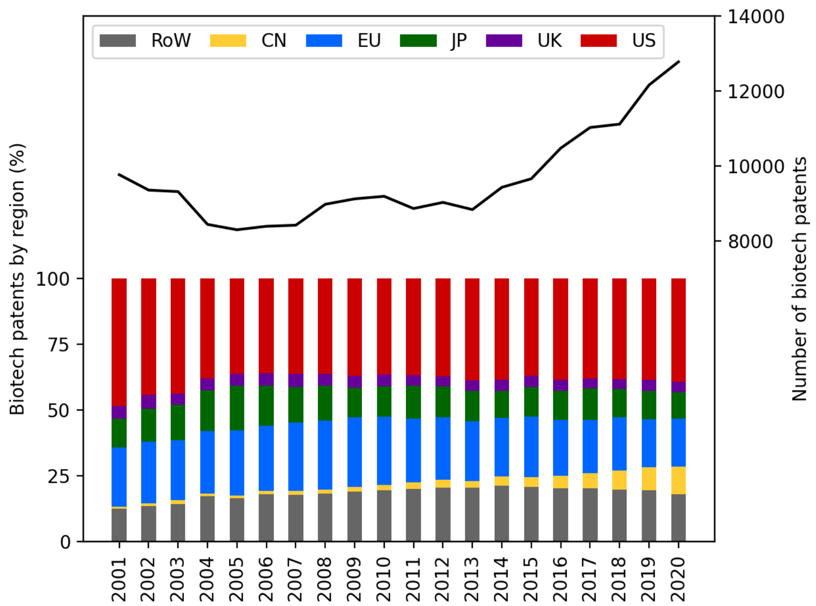
[10] See https://www.epo.org/en/news-events/in-focus/biotechnology-patents/red-white-green.

3 Trends in biotechnology patents

The bar chart at the bottom of figure 3 breaks down biotechnology patent filings geographically by attributing patents to the country or region (China, EU, Japan, UK, US, Rest of the World[11]) in which the applicants were located at the time of filing. According to the chart, the US (red) is the clear frontrunner in terms of the number of biotechnology patents, having a share of 48.5% of total biotechnology patents in 2001 that gradually reduced to just under 39.2% in 2020. The EU (blue) owns less than half of the biotechnology patents owned by the US (22.6% in 2001), but has followed a similar trend with its share gradually falling to 18.3% in 2020. Japan (green) and the UK (purple) follow with a smaller share (respectively, 10.0% and 4.1% in 2020), albeit more stable one. China started as the region with by far the lowest number of biotechnology patents but has grown noticeably, especially in the last decade, reaching a share of 10.4% of total biotechnology patents. Also the weight of countries grouped in the residual RoW category has grown during the period.

The black line chart at the top of figure 3 represents the number of biotechnology patents filed yearly by applicants recorded in the Patstat database between 2001 and 2020. The time series shows a steady growth starting from 2013, while the decade before is characterised by a more erratic pattern of ups and downs. However, an increase in the number of biotechnology patents from around 9700 in 2001 to over 12500 in 2020 can be observed, corresponding to a ~30% increase. This increase is substantial but also consistent with the global growth in patent filings. In fact, biotechnology patents account for a relatively stable share of around 5% of total IP5 families recorded over the analysed time window.

Figure 3. Trends in biotechnology patents



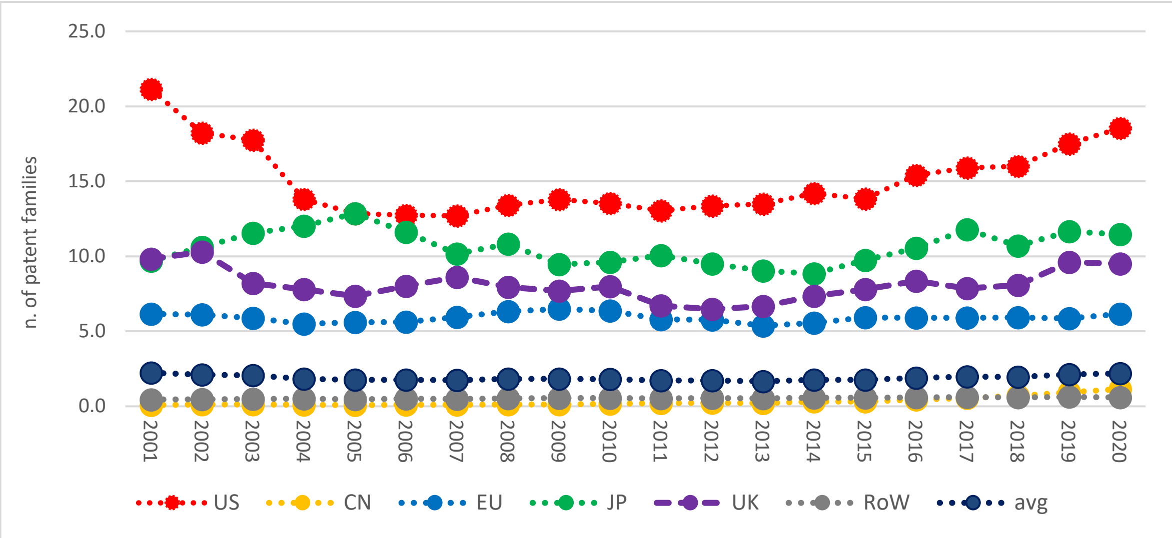
Note: Black line is a time series of the number of biotechnology patents filed worldwide between 2001 and 2020.
Bar chart: Share of biotechnology patents filed yearly by applicants located in each region.
Source: JRC computation based on Patstat 2023 Autumn Edition and the OECD MSTI statistics on patents by specific technology.

Figure 4 presents the same data but normalized by the population (population ages 15 and above). The role of the US as an innovation leader in the biotechnology field emerges clearly when looking at the data per capita. The gap between the EU and the US seems to decrease in the first decade and increase in the next one

[11] The choice of which countries (apart the EU) to include in the geographical disaggregation of the data is based on the ranking by number of biotechnology patents in the period 2001-2020. The top two countries of this ranking are the US and Japan. To this, China has been added, given its general global relevance and its specific growth in the field of biotechnology patents (especially in the last decade). All the other countries but the UK have been grouped in the residual category “Rest of the world”. This choice of presenting the UK not in the residual category is motivated by the fact that for 19 out of the 20 years analysed in the report the UK was part of the EU.

of the period considered. In 2001, the US had 3.4 times the number of patents per million inhabitants than the EU. This ratio declined to a minimum value of 2.1 in the years between 2007 and 2010 and then started to grow again reaching 3.0 in 2020. In the same period, the ratio between the EU and China in terms of biotechnology patents per million inhabitants fell considerably (from 73.5 to 5.3). This confirms that China has become an important player in the biotechnology patent field. Finally, it is interesting to notice how the relative importance of Japan and the UK increases when looking at the data per capita instead of the absolute numbers.

Figure 4. Trends in biotechnology patents per million of inhabitant

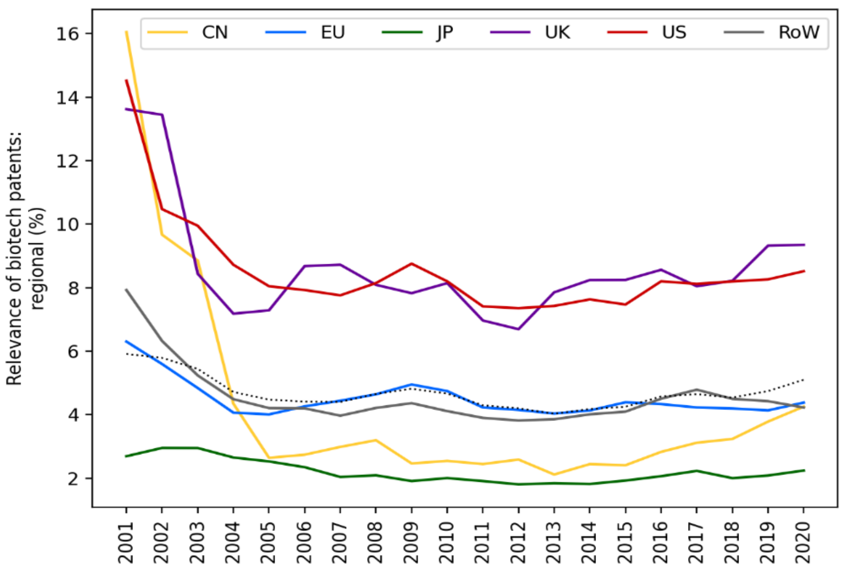


Note: Population ages 15 and above between 2001 and 2020.

Source: JRC computation based on Patstat 2023 Autumn Edition, the OECD MSTI statistics on patents by specific technology and the World Bank population data.

Figure 5 breaks this down further by region by plotting the evolution over time of the weight of biotechnology patents within the portfolio of each region. The line charts show a substantial heterogeneity across regions when it comes to the share of patents relating to biotechnology.

Figure 5. Weight of biotechnology patents in the regional/national patent portfolios

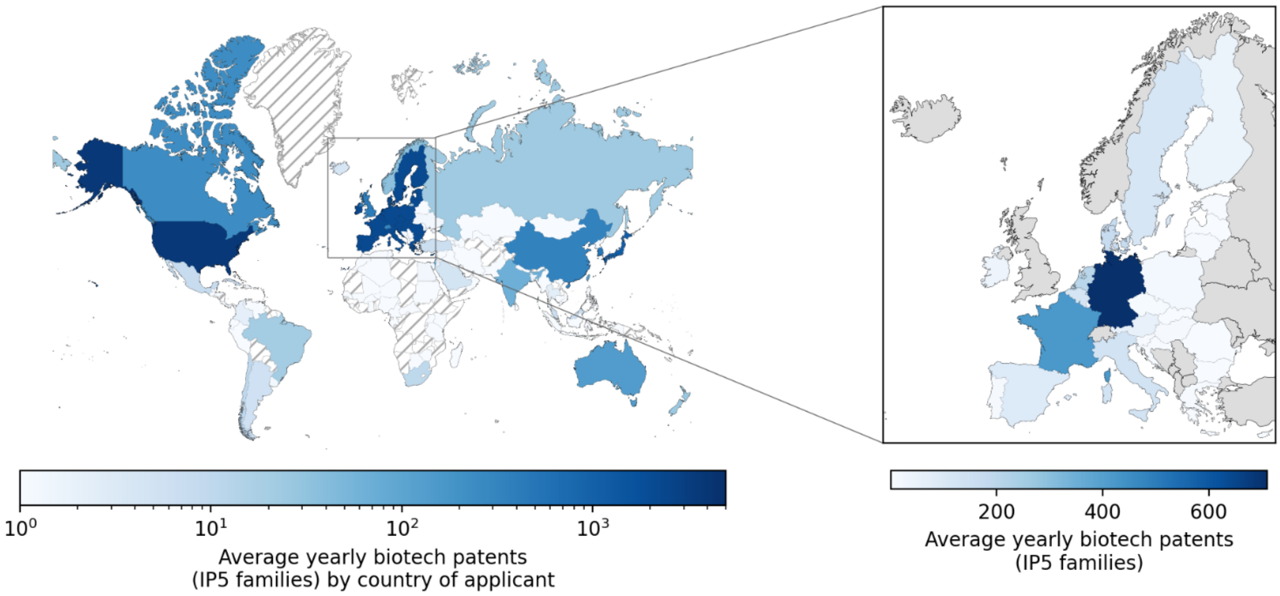


Source: JRC computation based on Patstat 2023 Autumn Edition and the OECD MSTI statistics on patents by specific technology.

In fact, in the US and the UK biotechnology patents represent consistently a higher share of the global patent portfolio than the abovementioned average of 5%. On the other hand, the EU and the RoW keep close to the global average throughout the time window, while Japan and China remain mostly below average. However, contrary to Japan, China starts converging toward the EU figure from 2015 onwards.

Figure 6 breaks down the bar chart of Figure 3 by reporting the average number of yearly biotechnology patents by applicant country between 2001 and 2020. Darker shades of blue represent larger values; the hatched pattern in the map identifies countires for which no data is available (i.e. countries that do not host applicants patenting in biotechnology according to the information available in the Patstat database). The world map on the left confirms the leading role of the US, followed by the EU, which, contrary to all the other territories in this panel, is reperesented as a unique region. The inset on the right zooms onto the EU, which, in this panel, is divided into its constituent Member States. This shows that there is a noticeable heterogeneity in the region, where Germany and France are responsible for most of the biotechnology patents, while eastern Members States are considerably less active in this specific field.

Figure 6. Average number of biotechnology patents filed worldwide between 2001 and 2020.

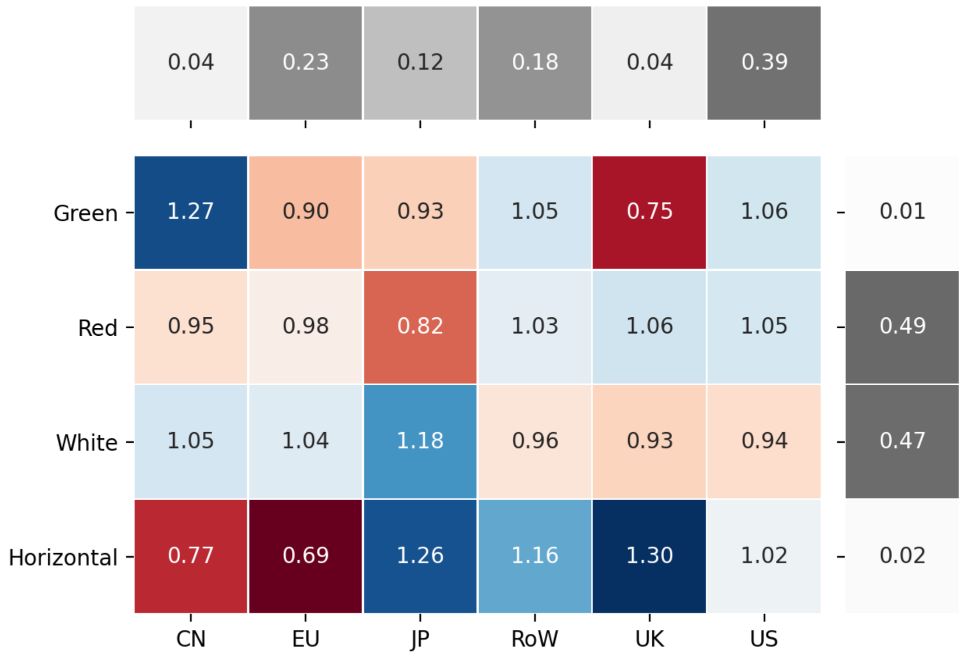


Note: Patents are allocated to the region of the applicant.
Source: JRC computation based on Patstat 2023 Autumn Edition and the OECD MSTI statistics on patents by specific technology.

4 Specialisation in biotech-related innovation

Figure 7 decomposes biotechnology patents by applicant region and *colour* categories (red, green, white, and horizontal) to compute a measure of relative specialisation in biotech-related innovation. Relative specialisation is computed by using the revealed comparative adavantage of regions in each biotechnology category; the values are reported in the large heatmap at the bottom left. The index tells us whether biotechnology patents of a given color weigh more or less within the patent portfolio of each region than biotechnology patents weigh globally relative to all IP5 patents. In a nutshell, a value of the index higher than one indicates that applicants in that region produce more than their fair share of biotechnology patents of a certain color. The coloured heat-map of Figure 7 reports the value of the index of relative technological specialization in biotechnology for the countries/regions analysed in the report. Blue coloured cells indicate combinations of combinations country/region and biotechnology area for which the Revealed Technological Advantage is greater than one; darker shades of blue stand for higher values of the index. Red coloured cells indicate the biotechnology areas in which regions have a Revealed Technological Advantage lower than one; darker shades of red stand for lower values of the index. The figure shows that Japanese, Chinese, and (to a lesser extent) EU applicants are relatively specialised in white biotechnology patents; UK applicants are relatively specialized in horizontal and, to a lesser extent, in red biotechnology. Chinese applicants are also relatively specialized in green biotechnology. The only field in which US applicants appear relatively less specialisedis white biotechnology. The greyscale heatmaps at the top and at the right of the figure help further interpret the colored heatmap. The heatmap on the right reports share of biotechnology patents belonging to each category. The heatmap at the top reports the share of biotechnology patents owned by each country/region. For insance, the US owns most of the biotechnology patents, while the UK owns by far the lowest share. At the same time, red and white biotechnology are by far the categories in which most patents are filed. Therefore, the relatively high specialization of the UK in horizontal biotechnology should be taken with a grain of salt, since horizontal biotechnology patents account for a relatively small share of biotechnology patents (2.1%) and the UK is, in general, responsible for only a small share of biotechnology patents (4.3%).

Figure 7. Relative technological specialization in biotech



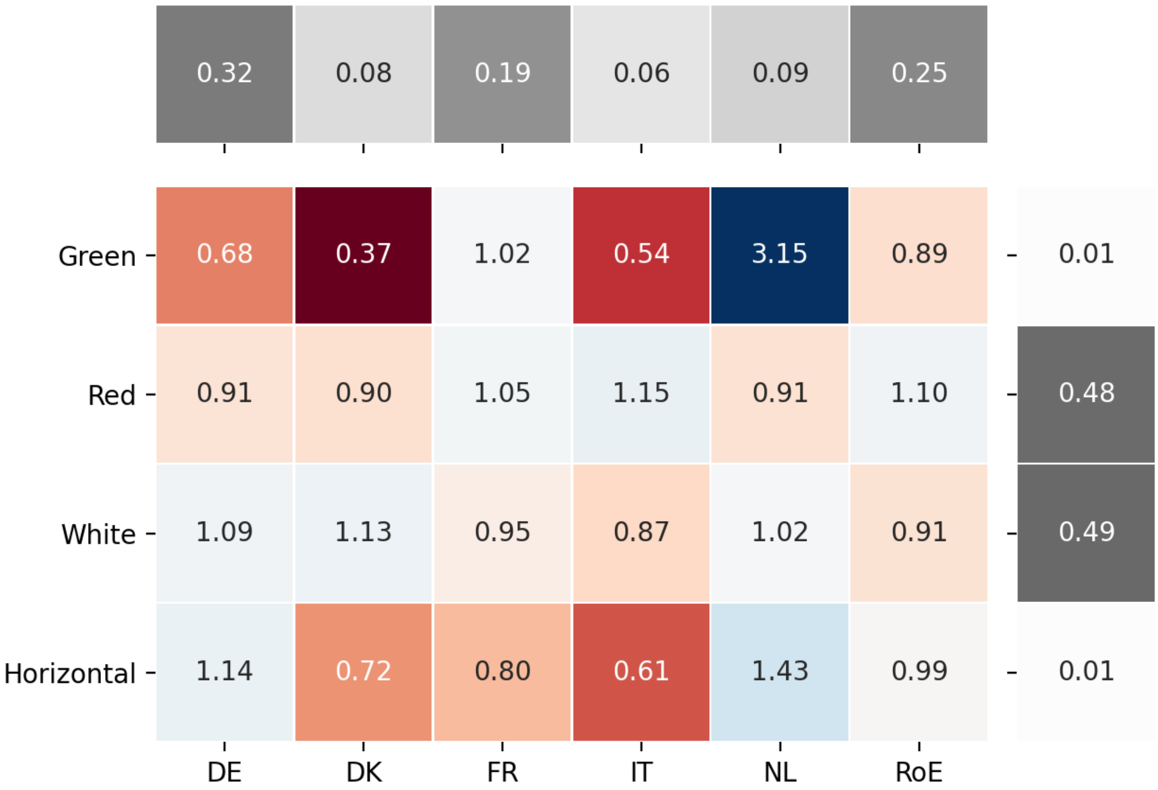
Source: JRC computation based on Patstat 2023 Autumn Edition and the OECD MSTI statistics on patents by specific technology.

Figure 8 replicates the analysis conducted in Figure 7, but specifically for the EU. It focuses on the five Member States with the highest share of biotechnology patents during the observed period, which accounted for 74.9% of the total EU biotechnology patents from 2001 to 2020.

Starting from the right column, which displays the relative shares of green, red, white, and horizontal biotechnologies filed by EU applicants, it is interesting to note that the distribution for EU applicants is almost the same as the overall distribution worldwide. The number of green biotechnology patents is extremely low, while white and red biotechnology groups represent the majority of all biotechnology patents. Unsurprisingly, when looking at the top column, Germany and France are the countries with the highest number of biotechnology patent applicants. Together, these two countries account for slightly more than 50% of all EU biotechnology patents.

Lastly, focusing on the relative advantage panel, the Netherlands is the only country that shows clear specialization in green biotechnology. Italy has the highest specialization index in red biotechnology, and Denmark has the highest index values for white biotechnology.

Figure 8. Relative technological specialization in biotechnology – EU focus

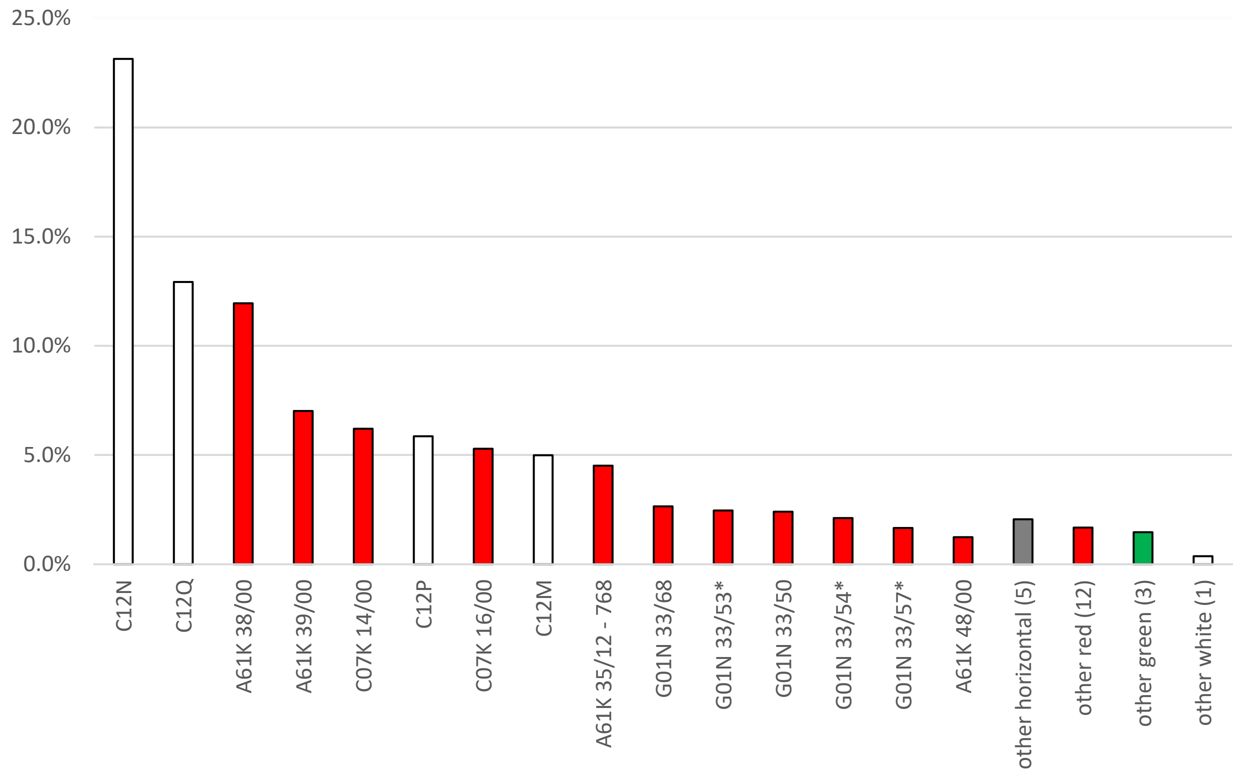


Source: JRC computation based on Patstat 2023 Autumn Edition and the OECD MSTI statistics on patents by specific technology.

5 A deeper look into biotechnology patents

Figure 9[12] delves deeper into the types of biotechnology that were patented during the observed period. A significant portion (23.1%) of all biotechnology patent families filed between 2001 and 2020 fall under the *C12N* category, which encompasses *"Micro-organisms or enzymes; compositions thereof (...); propagating, preserving, or maintaining micro-organisms; mutation or genetic engineering; culture media (microbiological testing media)"*. The second most prominent technological group is C12Q, accounting for 12.9% of all biotechnology patents. C12Q refers to *"Measuring or testing processes involving enzymes or micro-organisms (immunoassay); compositions or test papers therefor; processes of preparing such compositions; condition-responsive control in microbiological or enzymological processes"*. Both of these technologies are classified as white biotechnology and collectively account for over a third (36.1%) of all biotechnology patents filed worldwide during the observation period. The third most patented biotechnology is A61K 38/00, categorized as red biotechnology. It pertains to *"Medicinal preparations containing peptides"*. This category accounts for 11.9% of all biotechnology patents filed.

Figure 9. Top 15 biotechnology groups by share of patents worldwide 2001-2020



Note: Bars coloured according to the biotechnological group.
Source: JRC computation based on Patstat 2023 Autumn Edition and the OECD MSTI statistics on patents by specific technology.

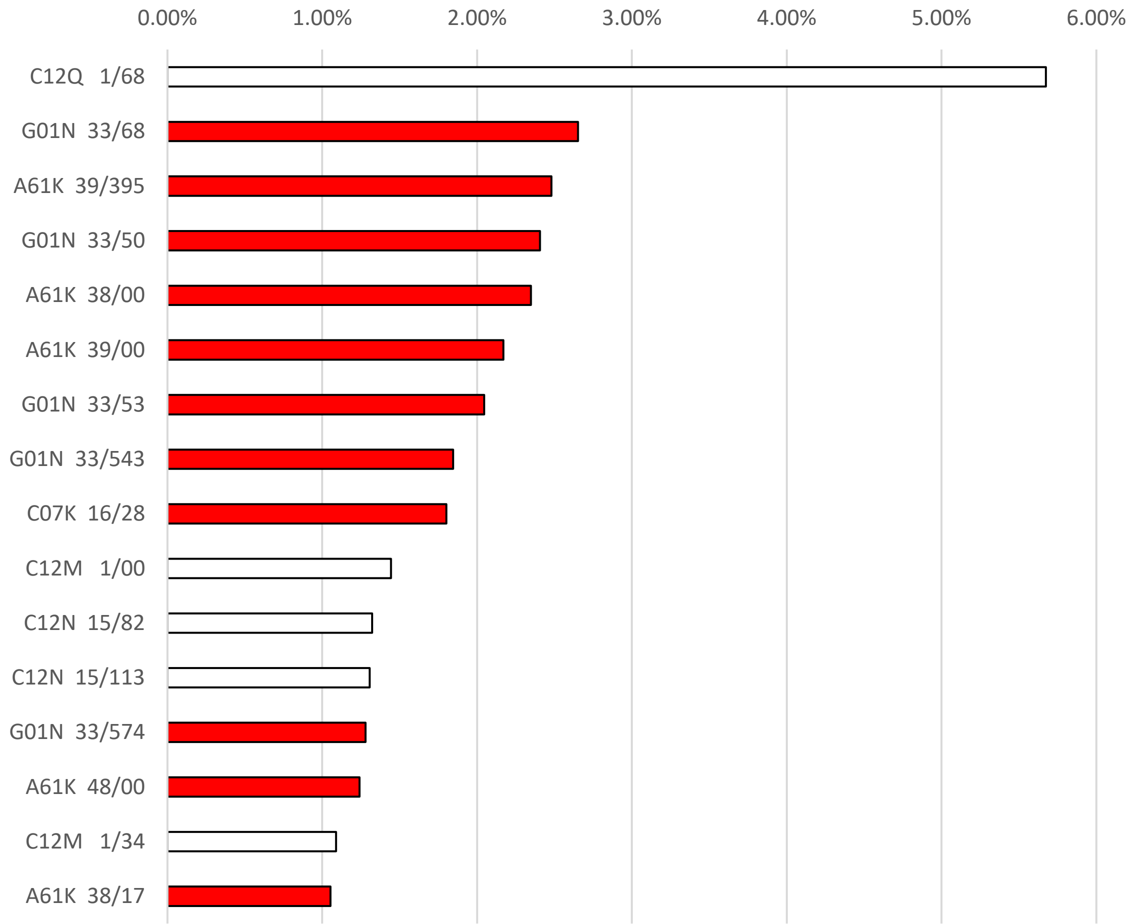
Figure 10 zooms in on Figure 9 and examines the 38 biotechnology groups previously analysed to identify the most relevant individual biotechnologies patented worldwide. Figure 10 presents the 16 individual technologies (out of 965) that represent at least 1% of the total biotechnology patents filed during the observed period[13]. Following the above, it is not surprising that the top technologies are predominantly classified as white (5) or red (11).

The technology that is by far the most patented is C12Q 1/68, specifically *" Measuring or testing processes involving nucleic acids"*, which alone represents 5.7% of all the biotechnology patents analysed. Completing the podium are two red biotechnologies: G01N 33/68, *Investigating or analysing materials by specific*

[12] See annex 1 for the correspondence between IPC codes and technology full names.
[13] See annex 2 for the full correspondence between IPC codes and technology names.

methods not covered by the preceding groups: involving proteins, peptides or amino acids", and A61K 39/395, *"Antibodies (agglutinins A61K 38/36); Immunoglobulins; Immune serum, e.g. antilymphocytic serum"*

Figure 10. Top biotechnology subclasses worldwide 2001-2020

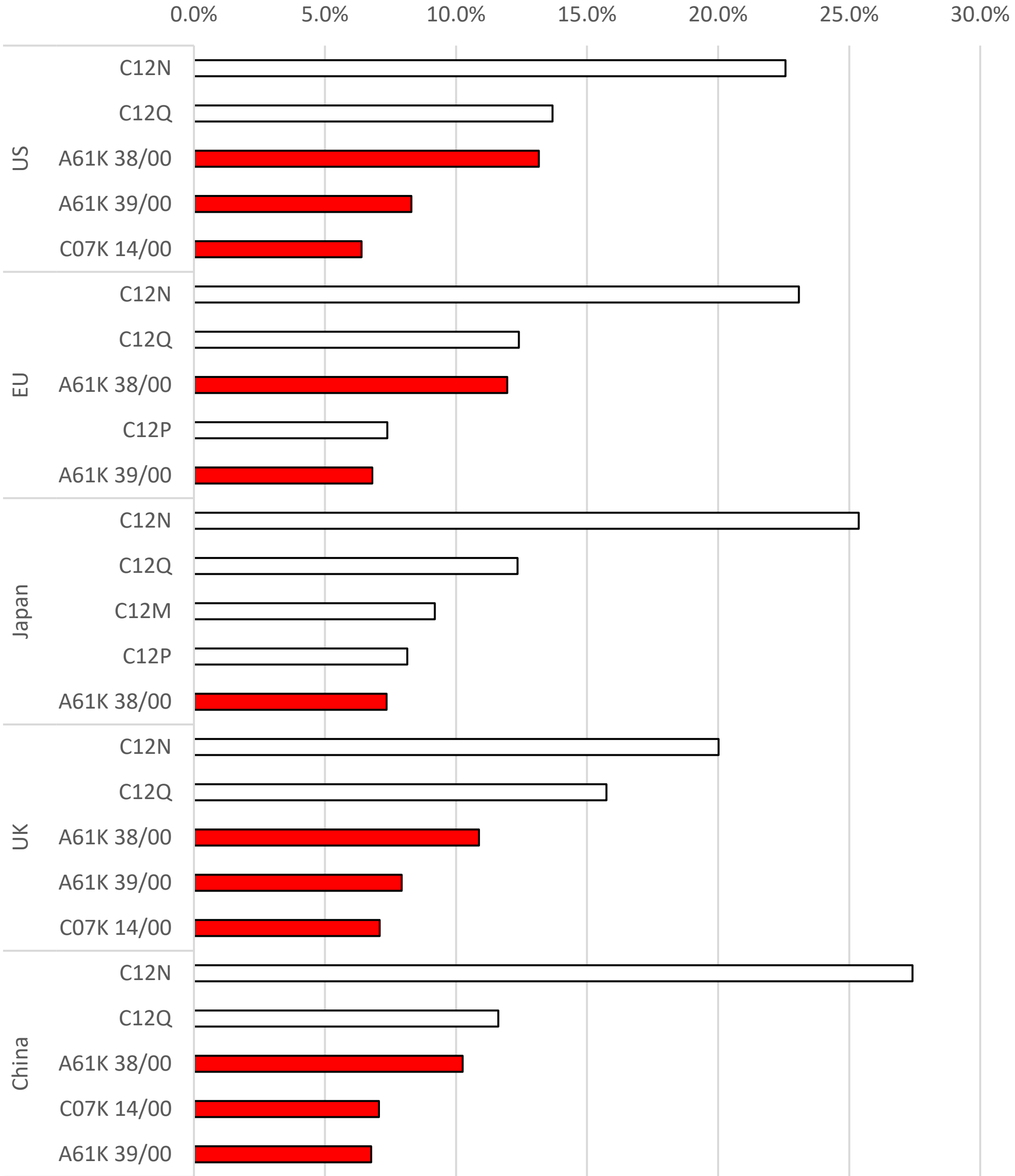


Note: Top biotechnology subclasses (representing each at least 1% of all the IP5 biotechnology patents filed worldwide 2001-2020).
Bars coloured according to the biotechnological group.
Source: JRC computation based on Patstat 2023 Autumn Edition and the OECD MSTI statistics on patents by specific technology.

Figure 11 also zooms into Figure 9 and compares the biotechnological profiles of the EU, US, Japan, UK and China. It displays the top 5 biotechnology classes based on the share of patents for each region. The first 3 are the same for all countries/regions (with the partial exception of Japan), aligning with what was seen in Figure 9 for the entire sample.

This seems to suggest that the competition among major players in biotechnology patents revolves around the number of patents in each of the main biotechnological domains rather than the different types of biotechnologies patented.

Figure 11. Top 5 biotechnology groups by share of patents per region 2001-2020.



Note: Bars coloured according to the biotechnological group.
Source: JRC computation based on Patstat 2023 Autumn Edition and the OECD MSTI statistics on patents by specific technology.

6 Conclusions

This document describes innovation in biotechnologies (*biotech*) as captured by patented inventions worldwide. To this aim, the report focuses on international patents filed at multiple offices, at least two of which belonging to the IP5 consortium. Moreover, it relies on expert knowledge collected by the OECD to select the inventions connected to biotechnology with a view to the evolution of patenting in this technological area over time and its relevance across geographical and technological dimensions.

The analysis presented in this report shows how relevant biotechnology patents are (around 5% of all the IP5 patents in the period 2001-2020). The US are by far the country with the highest share of biotechnology patents, followed by the EU (however with an increasing gap to the US), while China is catching up with the EU. The majority of the biotechnology patents are white (industrial) and red (medical) biotechnologies.

Japanese, Chinese, and (to a lesser extent) EU applicants are relatively specialised in white biotechnology patents. Chinese applicants are also relatively specialized in green biotechnology, while UK applicants are relatively specialized in horizontal and in red biotechnology. The only field in which US applicants appear relatively less specialised is white biotech. Germany and France have the highest number of biotechnology patent applicants in the EU, accounting for slightly over 50% of all EU biotechnology patents. The single biotechnology most patented is C12Q 1/68, *"Measuring or testing processes involving nucleic acids*", which alone represents 5.7% of all the biotechnology patents analysed.

Preliminary results suggest that the competition among countries/regions in biotechnology patents revolves around the number of patents in each of the main biotechnological domains, rather than the different types of biotechnologies patented.

References

Amoroso S., Aristodemou L., Criscuolo C., Dechezleprêtre A., Dernis H., Grassano N., Moussiegt L., Napolitano L., Nawa D. Squicciarini M., Tübke A. (2021). World Corporate Top R&D investors: Paving the way for climate neutrality. A joint JRC and OECD report. EUR 30884 EN, Publications Office of the European Union, Luxembourg, 2021, ISBN 978-92-76-43373-6, doi:10.2760/49552, JRC126788.

Friedrichs, S. and B. van Beuzekom (2018), "Revised proposal for the revision of the statistical definitions of biotechnology and nanotechnology", *OECD Science, Technology and Industry Working Papers*, No. 2018/01, OECD Publishing, Paris, https://doi.org/10.1787/085e0151-en.

Grassano, N., Hernandez Guevara, H., Fako, P., Tübke, A., Amoroso, S., Georgakaki, A., Napolitano, L., Pasimeni, F., Rentocchini, F., Compaño, R., Fatica, S. and Panzica, R. The 2021 EU Industrial R&D Investment Scoreboard, EUR 30902 EN, Publications Office of the European Union, Luxembourg, 2022, ISBN 978-92-76-44399-5, doi:10.2760/559391, JRC127360.

Mueller, D. C. (1966). Patents, research and development, and the measurement of inventive activity. *The Journal of Industrial Economics*, 26-37.

Pavitt, K. (1985). Patent statistics as indicators of innovative activities: possibilities and problems. *Scientometrics*, *7*(1-2), 77-99.

Ronzon, T., Iost, S., & Philippidis, G. (2022). An output-based measurement of EU bioeconomy services: Marrying statistics with policy insight. Structural Change and Economic Dynamics, 60, 290-301. https://doi.org/10.1016/j.strueco.2021.10.005.

List of abbreviations and definitions

CNIPA	China National Intellectual Property Administration
EPO	European Patent Office
IP5	Consortium of five patent offices (EPO, USPTO, JPO, KIPO, CNIPA)
JPO	Japanese Patent Office
KIPO	Korean Intellectual Property Office
USPTO	United States Patent and Trademark Office

List of figures

Annexes

Annex 1. List of biotechnology IPC codes.

The first 2 columns of the table below are taken from the "Revised proposal for the revision of the statistical definitions of biotechnology and nanotechnology" published online by the OECD (available at https://doi.org/10.1787/085e0151-en). The colour coding of the different technologies (third column) is the result of internal analysis.

IPC Symbol	Title	Colour
A01H 1/00	Processes for modifying genotypes	Green
A01H 4/00	Plant reproduction by tissue culture techniques	Green
A01K 67/00	Rearing or breeding animals, not otherwise provided for; New breeds of animals	Green
A61K 35/12 - 768	Materials from mammals; Compositions comprising non-specified tissues or cells; Compositions comprising non-embryonic stem cells; Genetically modified cells (uncharacterised stem cells; vaccines or medicinal preparations containing antigens or antibodies)	Red
A61K 38/00	Medicinal preparations containing peptides (peptides containing beta-lactam rings; cyclic dipeptides not having in their molecule any other peptide link than those which form their ring, e.g. piperazine-2,5-diones; ergoline-based peptides; containing macromolecular compounds having statistically distributed amino acid units; medicinal preparations containing antigens or antibodies; medicinal preparations characterised by the non-active ingredients, e.g. peptides as drug carriers)	Red
A61K 39/00	Medicinal preparations containing antigens or antibodies (materials for immunoassay)	Red
A61K 48/00	Medicinal preparations containing genetic material which is inserted into cells of the living body to treat genetic diseases; Gene therapy	Red
C02F 3/34	Biological treatment of water, waste water, or sewage: **characterised by the micro-organisms used**	White
C07G 11/00	Organic chemistry - compounds of unknown constitution: **antibiotics**	Red
C07G 13/00	Organic chemistry - compounds of unknown constitution: **vitamins**	Red
C07G 15/00	Organic chemistry - compounds of unknown constitution: **hormones**	Red
C07K 4/00	Organic chemistry - peptides having up to 20 amino acids in an undefined or only partially defined sequence; Derivatives thereof	Red
C07K 14/00	Organic chemistry - peptides having more than 20 amino acids; Gastrins; Somatostatins; Melanotropins; Derivatives thereof	Red
C07K 16/00	Organic chemistry - Immunoglobulins, e.g. monoclonal or polyclonal antibodies.	Red
C07K 17/00	Organic chemistry - Carrier-bound or immobilised peptides; Preparation thereof	Red
C07K 19/00	Organic chemistry - Hybrid peptides (hybrid immunoglobulins composed solely of immunoglobulins)	Red
C12M	APPARATUS FOR ENZYMOLOGY OR MICROBIOLOGY (installations for fermenting manure; preservation of living parts of humans or animals; brewing apparatus; fermentation apparatus for wine; apparatus for preparing vinegar)	White
C12N	MICRO-ORGANISMS OR ENZYMES; COMPOSITIONS THEREOF (biocides, pest repellants or attractants, or plant growth regulators containing micro-organisms, viruses, microbial fungi, enzymes, fermentates, or substances produced by, or extracted from, micro-organisms or animal material; medicinal preparations; fertilisers); PROPAGATING, PRESERVING, OR MAINTAINING MICRO-ORGANISMS; MUTATION OR GENETIC ENGINEERING; CULTURE MEDIA (microbiological testing media)	White

C12P	FERMENTATION OR ENZYME-USING PROCESSES TO SYNTHESISE A DESIRED CHEMICAL COMPOUND OR COMPOSITION OR TO SEPARATE OPTICAL ISOMERS FROM A RACEMIC MIXTURE	White
C12Q	MEASURING OR TESTING PROCESSES INVOLVING ENZYMES OR MICRO-ORGANISMS (immunoassay); COMPOSITIONS OR TEST PAPERS THEREFOR; PROCESSES OF PREPARING SUCH COMPOSITIONS; CONDITION-RESPONSIVE CONTROL IN MICROBIOLOGICAL OR ENZYMOLOGICAL PROCESSES	White
C40B 10/00	Directed molecular evolution of macromolecules, e.g. RNA, DNA or proteins	horizontal
C40B 40/02	Libraries per se, e.g. arrays, mixtures: **Libraries contained in or displayed by micro-organisms or vectors**	horizontal
C40B 40/06	Libraries per se, e.g. arrays, mixtures: **Libraries containing nucleotides or polynucleotides, or derivatives thereof**	horizontal
C40B 40/08	Libraries per se, e.g. arrays, mixtures: **Libraries containing peptides or polypeptides, or derivatives thereof**	horizontal
C40B 50/06	Methods of creating libraries, e.g. combinatorial synthesis: **Biochemical methods, e.g. using enzymes or whole viable micro-organisms**	horizontal
G01N 27/327	Investigating or analysing materials by the use of electric, electro-chemical, or magnetic means: **biochemical electrodes**	horizontal
G01N 33/50	Investigating or analysing materials by specific methods not covered by groups: **Chemical analysis of biological material, e.g. blood, urine; Testing involving biospecific ligand binding methods; Immunological testing (measuring or testing processes other than immunological involving enzymes or micro-organisms, compositions or test papers therefor; processes of forming such compositions, condition responsive control in microbiological or enzymological processes**	Red
G01N 33/53*	Investigating or analysing materials by specific methods not covered by the preceding groups: **immunoassay; biospecific binding assay; materials therefor**	Red
G01N 33/54*	Investigating or analysing materials by specific methods not covered by the preceding groups: **double or second antibody: with steric inhibition or signal modification: with an insoluble carrier for immobilising immunochemicals: the carrier being organic: synthetic resin: as water suspendable particles**: with antigen or antibody attached to the carrier via a bridging agent: Carbohydrates: with antigen or antibody entrapped within the carrier	Red
G01N 33/55*	Investigating or analysing materials by specific methods not covered by the preceding groups: **the carrier being inorganic: Glass or silica: Metal or metal coated: the carrier being a biological cell or cell fragment: Red blood cell: Fixed or stabilised red blood cell: using kinetic measurement: using diffusion or migration of antigen or antibody: through a gel**	Red
G01N 33/57*	Investigating or analysing materials by specific methods not covered by the preceding groups: **for venereal disease: for enzymes or isoenzymes: for cancer: for hepatitis: involving monoclonal antibodies: involving limulus lysate**	Red
G01N 33/68	Investigating or analysing materials by specific methods not covered by the preceding groups: **involving proteins, peptides or amino acids**	Red
G01N 33/74	Investigating or analysing materials by specific methods not covered by the preceding groups: **involving hormones**	Red
G01N 33/76	Investigating or analysing materials by specific methods not covered by the preceding groups: **human chorionic gonadotropin**	Red

G01N 33/78	Investigating or analysing materials by specific methods not covered by the preceding groups: **thyroid gland hormones**	Red
G01N 33/88	Investigating or analysing materials by specific methods not covered by the preceding groups: **involving prostaglandins**	Red
G01N 33/92	Investigating or analysing materials by specific methods not covered by the preceding groups: **involving lipids, e.g. cholesterol**	Red
G06F 19/10 – 24	Digital computing or data processing equipment or methods, specially adapted for specific applications (data processing systems or methods specially adapted for administrative, commercial, financial, managerial, supervisory or forecasting purposes): **Bioinformatics, i.e. methods or systems for genetic or protein-related data processing in computational molecular biology (in silico methods of screening virtual chemical libraries; in silico or mathematical methods of creating virtual chemical libraries)**	horizontal

Annex 2. List of biotechnologies sub classes representing at least 1% of the total biotechnology IP5 patents filed in the period 2001-2020

IPC Symbol	Subcategory	Title
C12Q	C12Q 1/68	Measuring or testing processes involving nucleic acids
G01N 33/68	G01N 33/68	Investigating or analysing materials by specific methods not covered by the preceding groups: involving proteins, peptides or amino acids
A61K 39/00	A61K 39/395	Antibodies (agglutinins A61K 38/36); Immunoglobulins; Immune serum, e.g. antilymphocytic serum
G01N 33/50	G01N 33/50	Investigating or analysing materials by specific methods not covered by groups: Chemical analysis of biological material, e.g. blood, urine; Testing involving biospecific ligand binding methods; Immunological testing (measuring or testing processes other than immunological involving enzymes or micro-organisms, compositions or test papers therefor; processes of forming such compositions, condition responsive control in microbiological or enzymological processes
A61K 38/00	A61K 38/00	Medicinal preparations containing peptides (peptides containing beta-lactam rings A61K 31/00; cyclic dipeptides not having in their molecule any other peptide link than those which form their ring, e.g. piperazine-2,5-diones, A61K 31/00; ergoline-based peptides A61K 31/48; containing macromolecular compounds having statistically distributed amino acid units A61K 31/74; medicinal preparations containing antigens or antibodies A61K 39/00; medicinal preparations characterised by the non-active ingredients, e.g. peptides as drug carriers, A61K 47/00)
A61K 39/00	A61K 39/00	Medicinal preparations containing antigens or antibodies (materials for immunoassay G01N 33/53)
G01N 33/53*	G01N 33/53	Investigating or analysing materials by specific methods not covered by the preceding groups: Immunoassay; Biospecific binding assay; Materials therefor
G01N 33/54*	G01N 33/543	Investigating or analysing materials by specific methods not covered by the preceding groups: with an insoluble carrier for immobilising immunochemicals
C07K 16/00	C07K 16/28	Immunoglobulins, e.g. monoclonal or polyclonal antibodies: against receptors, cell surface antigens or cell surface determinants
C12M	C12M 1/00	Apparatus for enzymology or microbiology
C12N	C12N 15/82	Vectors or expression systems specially adapted for eukaryotic hosts for plant cells
C12N	C12N 15/113	Processes for the isolation, preparation or purification of DNA or RNA: Non-coding nucleic acids modulating the expression of genes, e.g. antisense oligonucleotides
G01N 33/57*	G01N 33/574	Investigating or analysing materials by specific methods not covered by the preceding groups: for cancer
A61K 48/00	A61K 48/00	Medicinal preparations containing genetic material which is inserted into cells of the living body to treat genetic diseases; Gene therapy
C12M	C12M 1/34	Measuring or testing with condition measuring or sensing means, e.g. colony counters
A61K 38/00	A61K 38/17	Medicinal preparations containing peptides: from animals; from humans