Literature review of the trends and issues in synthetic biology (2012–2023)

Austein McLoughlin

Programme Management Officer

Secretariat of the Convention on Biological Diversity

Overview

- Prepared by the Secretariat with the support of a consultant and the financial support of the European Union
- Peer review process during January and February 2024
- Based on trends and issues in provisional selection list (55 in total)
  - Operational definition of synthetic biology
  - 2012 to 2023
- Aims to be a complementary quantitative analysis for the broad and regular horizon scanning, monitoring and assessment process

See document CBD/SBSTTA/26/INF/5 for full detail

Methodology

Published literature analysis

Acquisition

Scopus®

- Boolean search strings (bucket term +

Data cleaning

- Search string adjustments
- Duplicate removal

Analyses

- Temporal indicators
- Geographic information
- Co-authorship collaboration

Patent analysis

Acquisition

WIPO PATENTSCOPE

- Boolean search string
- Specific classes

Data extraction

- Class and subgroup breakdown
- Geographical information

Analyses

- Temporal indicators
- Class and subgroup analysis
- Geographic information

Results – Published literature

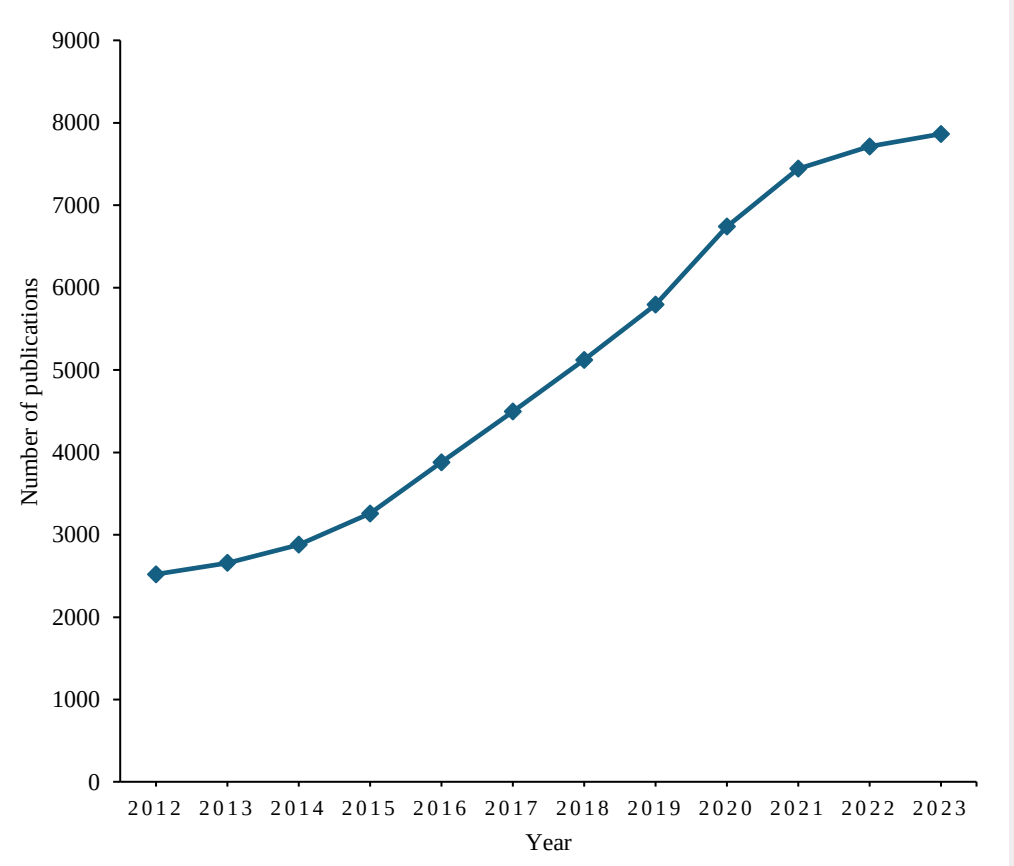
- Total of 60,358 publications for all 55 trends and issues
- Top 10:

	Identified issues and trends in synthetic biology	Publications	Percentage
1	Medical and therapeutic synthetic biology applications	19 797	32.80
2	Genome edited plants	7 896	13.08
3	Synthetic biology-enabled production of antibiotics, natural products and medically relevant compounds	6 745	11.17
4	Increased sophistication and expansion of genome editing tools	4 895	8.11
5	Synthetic biology-enabled production of petrochemical precursors and industrial chemicals	4 371	7.24
6	Advances in protein engineering	4 311	7.14
7	Integration of artificial intelligence and machine learning	3 950	6.54
8	Genome edited animals	2 691	4.46
9	Synthetic biology applications for bioremediation, biodegradation or biomining	2 513	4.16
10	Biosensors, sensory devices and diagnostics	2 043	3.38

Results – Published literature

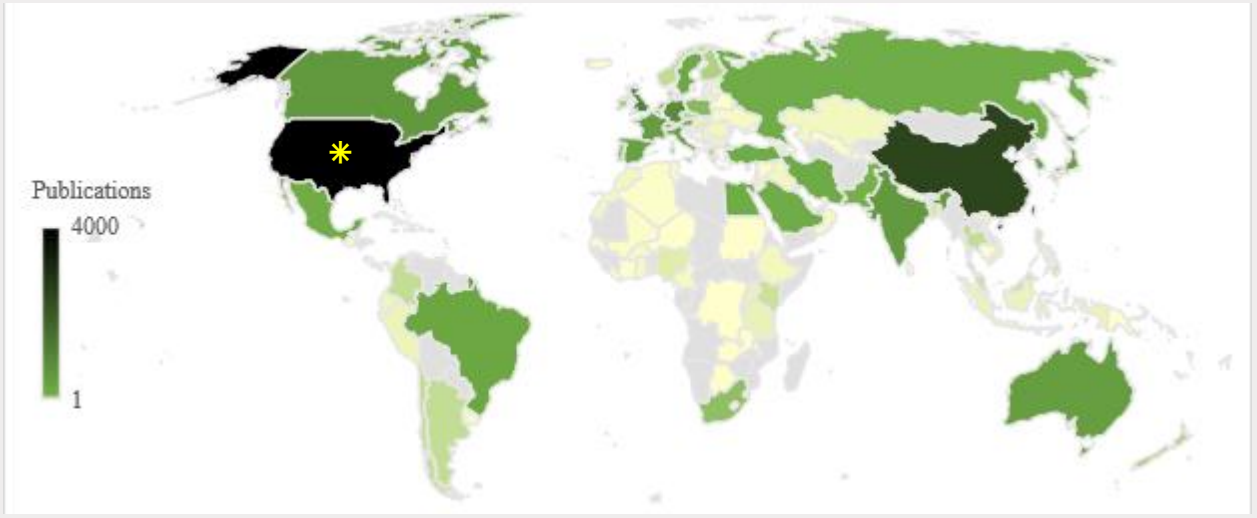
- Annual publication rate increased linearly
  - 2,500 to 7,900
- Publication rate for individual topics
  - Constant or increasing linearly
- Highest increases related to genome editing
- Some others with strong linear growth included:
  - Artificial intelligence and machine learning
  - Next-generation sequencing and bioinformatics
  - Non-medical microbiome engineering
- Constant publication rates:
  - Protein engineering
  - Synthetic biology-enable production of petrochemical precursors and industrial chemicals

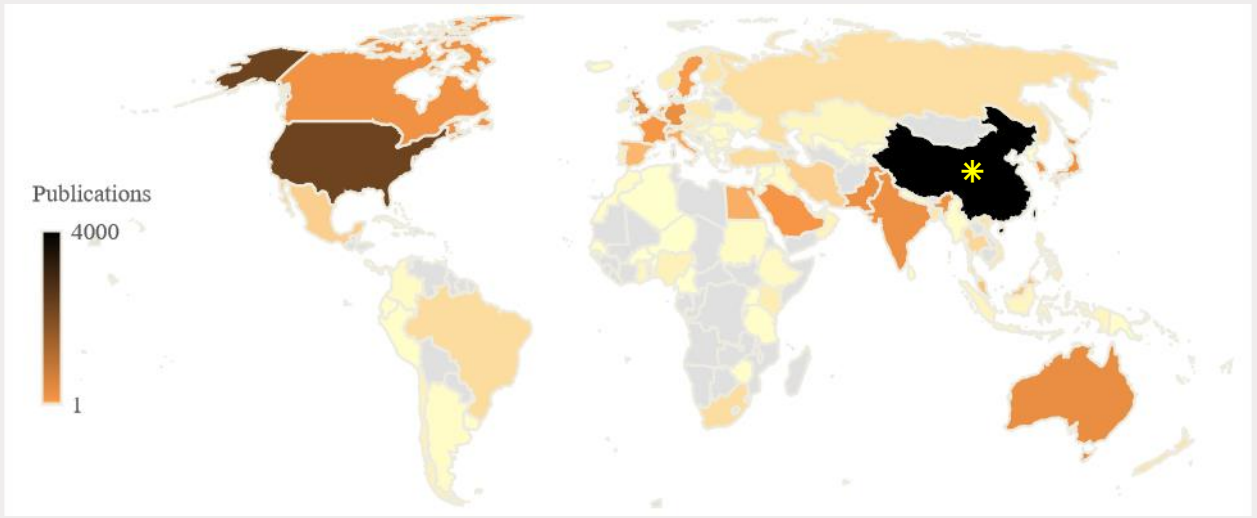
Publication rate of all 55 trends and issues



Results – Published literature

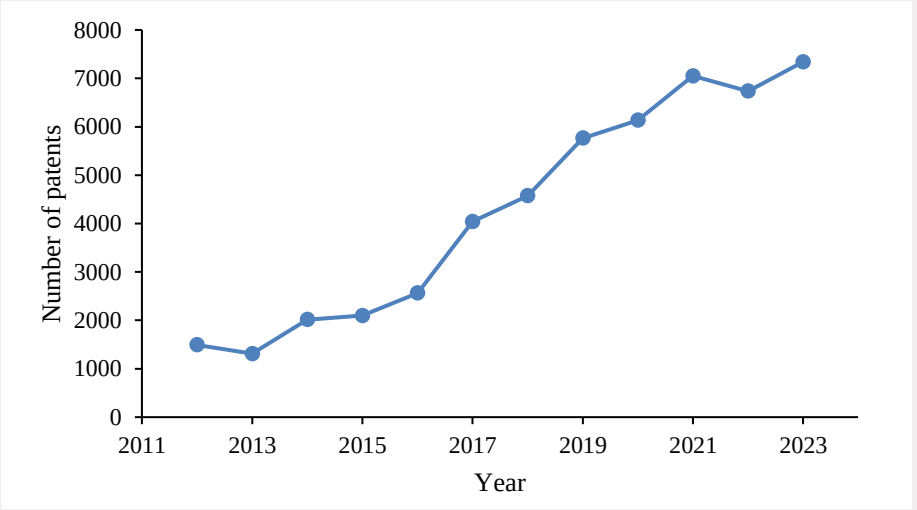
- The countries with the highest publication output:
  - United States of America
  - China
  - United Kingdom of Great Britain and Northern Ireland
  - India
  - Germany
- Co-authorship:
  - Patterns between lead authors and other researcher groups in other countries
  - Top 5 countries: at least 100 other countries
  - United States of America (green map)
  - China (orange map)

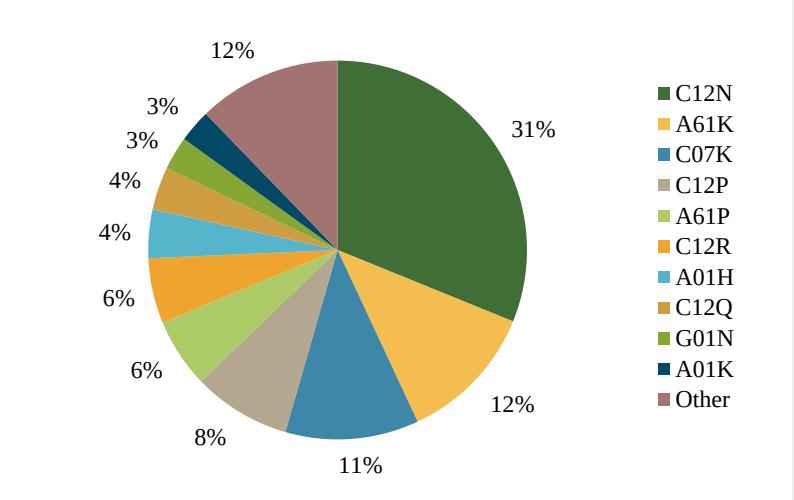




Results – Patent analysis – Bucket term

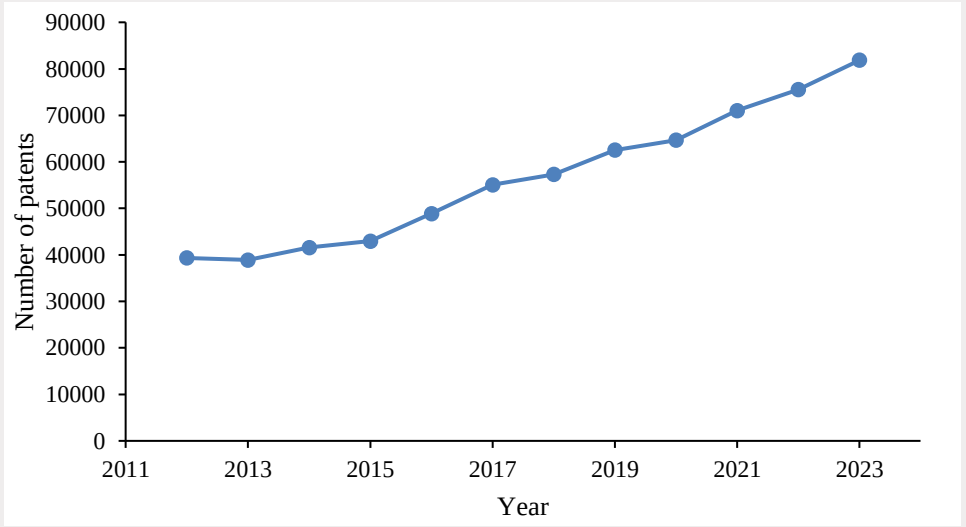
- Patent search of overall bucket term
  - 51,153 patents for 2012–2023
  - 1,492 to 7,343 between 2012–2023 (roughly linear growth in annual patent filings)
- Top classes:
  - C12N (Microorganisms or enzymes; compositions thereof; propagating, preserving, or maintaining microorganisms; mutation or genetic engineering; culture media)
  - A61K (Preparations for medical, dental or toilet purposes)
  - C07K (Peptides)
- Top countries:
  - China
  - United States of America
  - Australia
  - Canada
  - India

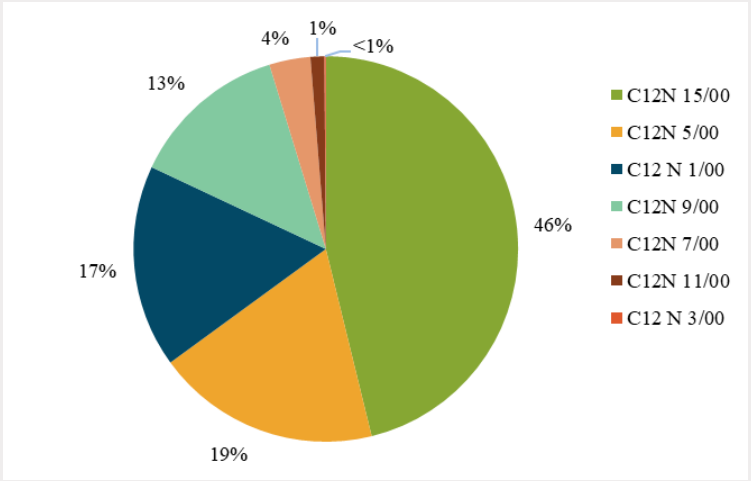




Results – Patent analysis – C12N[1]

- C12N
  - 679,807 patents for 2012–2023
  - 39,359 to 81,902 between 2012–2023 (near linear growth in annual patent filings)
- Top subgroups:
  - C12N 15/00 (Mutation or genetic engineering; DNA or RNA concerning genetic engineering, vectors, e.g., plasmids, or their isolation, preparation or purification; Use of hosts therefor)
  - C12N 5/00 (Undifferentiated human, animal or plant cells, e.g., cell lines; Tissues; Cultivation or maintenance thereof; Culture media therefor)
  - C12N 1/00 (Microorganisms, e.g., protozoa; Compositions thereof; Processes of propagating, maintaining or preserving microorganisms or compositions thereof; Processes of preparing or isolating a composition containing a microorganism; Culture media therefor)
- Top countries:
  - China
  - United States of America
  - Japan
  - Republic of Korea
  - Canada





[1]Microorganisms or enzymes; compositions thereof; propagating, preserving, or maintaining microorganisms; mutation or genetic engineering; culture media

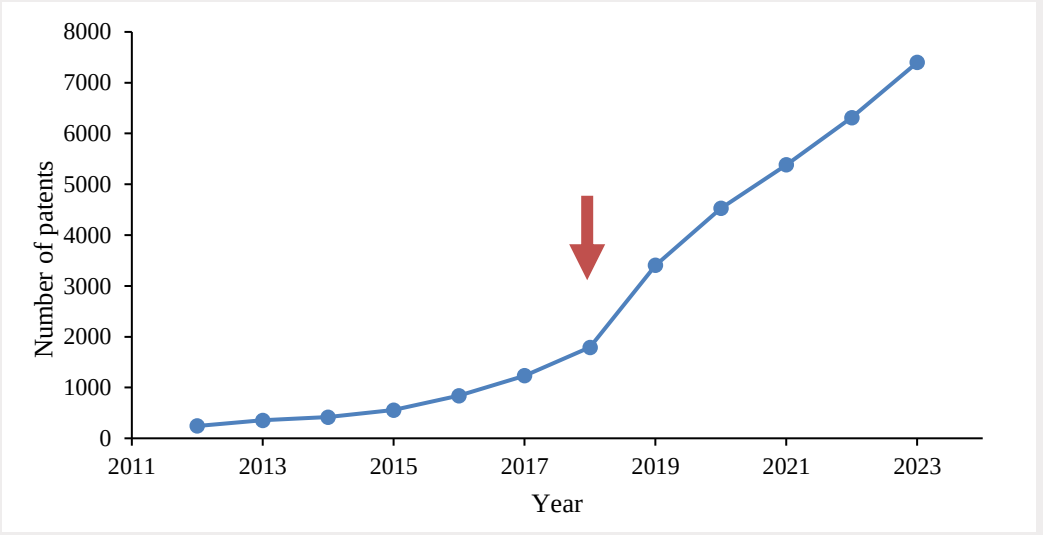
THE BIODIVERSITY PLAN
For Life on Earth

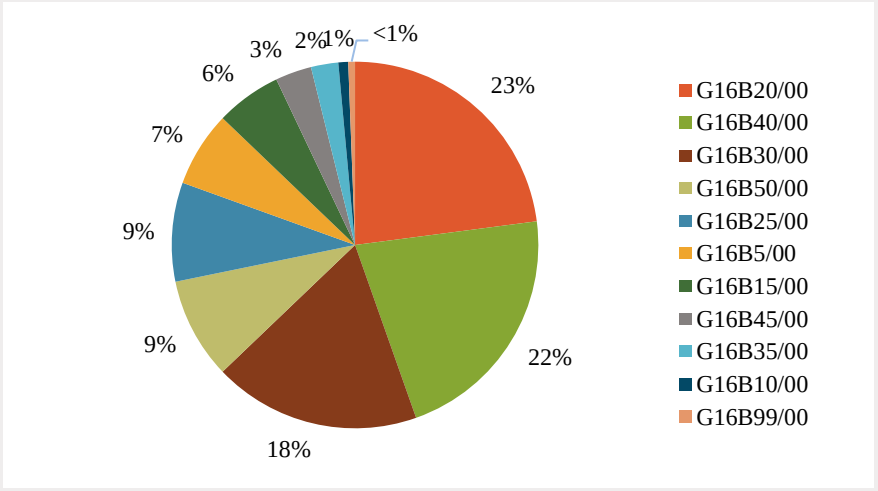
UN environment programme

Convention on Biological Diversity

Results – Patent analysis – G16B[2]

- G16B
  - 32,476 patents for 2012–2023
  - 246 to 7,401 annual patents filed between 2012–2023
- Top subgroups:
  - G16B 20/00 (functional genomics or proteomics, e.g., genotype-phenotype associations)
  - G16B 40/00 (biostatistics; bioinformatics-related machine learning or data mining, e.g., knowledge discovery or pattern finding)
  - G16B 30/00 (sequence analysis involving nucleotides or amino acids)
- Top countries:
  - China
  - United States of America
  - Republic of Korea
  - Canada
  - Japan





THE BIODIVERSITY PLAN
For Life on Earth

[2]Bioinformatics, i.e., information and communication technology [ICT] specially adapted for genetic or protein-related data processing in computational molecular biology

UN environment programme

Convention on Biological Diversity

Limitations

- Time
  - Conducted in January 2024, redone in March-April 2024
  - Asynchronous completion in relation to multidisciplinary AHTEG
- Methodology
  - Limited to Scopus in English
  - Based on Boolean searches (how the authors self-describe their work)
  - Limited data cleaning
  - Additional statistics and quantitative measures
- Trends and issues
  - Overlaps between trends and issues
  - Not all trends and issues can be measured using a bibliometric approach
- Current patent classification system
  - No differentiation between biotechnology and synthetic biology

Discussion and conclusions

- Overall linear increase in the annual rate of publications and patent filings
- Synthetic biology-enable production topics had moderate growth
- Similar patterns can be observed between the topics in the publications and the patent filings
  - Medical and therapeutic applications
  - Bioinformatics + artificial intelligence and machine learning
  - Protein engineering
- Contribution to publications vs patents
- Collaboration opportunities exist between countries
  - Concentration in developed countries

Thank you!

Secretariat of the Convention on Biological Diversity

413 St. Jacques Street, Suite 800
Montreal, Quebec, Canada H2Y 1N9
Tel. +1 514 288 2220

secretariat@cbd.int
bch@cbd.int

cbd.int
bch.cbd.int
bch.cbd.int/protocol

facebook.com/UNBiodiversity
facebook.com/UN.Biosafety

twitter.com/UNBiodiversity
twitter.com/BCHCPB

instagram.com/UNBiodiversity

linkedin.com/company/UNBiodiversity

Comparison to expert-driven process

Trend or issue	Multidisciplinary AHTEG process	Literature analysis
Self-spreading vaccines for wildlife	Prioritized for detailed assessment	**46th**
Self-limiting insect systems	Prioritized for detailed asssessment	**33rd**
Development of engineered gene drives to control vector-borne diseases and invasive species	Prioritized for detailed assessment	**23rd**
Integration of artificial intelligence and machine learning	Prioritized for detailed asssessment	**7th**
Inequity in the participation of developing countries in the context of synthetic biology	Prioritized for detailed assessment	**49th**
Medical and therapeutic synthetic biology applications	Not shortlisted	**1st**
Genome edited plants	Prioritized list (additional 12)	**2nd**
Synthetic biology-enabled production of antibiotics, natural products and medically relevant compounds	Shortlisted, not prioritized	**3rd**
Increased sophistication and expansion of genome editing tools	Shortlisted, not prioritized	**4th**
Synthetic biology-enabled production of petrochemical precursors and industrial chemicals	Shortlisted, not prioritized	**5th**

Methodology

Published literature analysis

Acquisition
- Scopus
- Boolean search strings (bucket term + specific topic)
- 55 trends + issues
- 2012 to 2023

→ **Inclusion criteria**
- Articles, conference proceedings, reviews, book chapters, surveys, editorials and notes
- English

→ **Data cleaning**
- Search string adjustments
- Duplicate removal

→ **Analyses**
- Temporal indicators
- Geographic information
- Co-authorship collaboration

Patent analysis

Acquisition
- PATENTSCOPE
- Boolean search string
- Specific classes
- 2012 to 2023

→ **Data extraction**
- Class and subgroup breakdown
- Geographical information

→ **Analyses**
- Temporal indicators
- Class and subgroup analysis
- Geographic information