



## Biotechnology in Developing Countries: Growth and Competitiveness

November 2016



**CAS-TWAS**  
Centre of Excellence  
for Biotechnology

**Clarivate  
Analytics**

Formerly the IP & Science  
business of Thomson Reuters

## Foreword

Biotechnology is an area of immense strategic importance as the world strives to adapt and respond to growing population, a shortage of resources and environmental pollution. Because of its many highly useful applications, biotech is becoming a major driver of economic growth and a key area of global competition.

Presently, there is a worrisome global gap in terms of biotechnology research and development (R&D) and development of related industries. Developed nations have taken the lead in this field, thanks to their advantages in human skill and technology; developing countries lag much behind. This unfortunate gap results both from lack of funds and human capital and from inadequate attention on the part of governments and policymakers.

Embracing the great opportunities of the emerging bio-economy, developing countries need to sharpen their awareness, increase their commitment and gradually build their own strength in the dynamic fields of biotechnology. Only in this way can they fully exploit its potential to drive economic growth and social development.

To achieve such progress, it is crucial to have a clear understanding of the current status of scientific research and productivity in all the fields of biotechnology.

The CAS-TWAS Centre of Excellence for Biotechnology (CoEBio), which was co-founded by the Chinese Academy of Sciences (CAS) and The World Academy of Sciences for the advancement of science in developing countries (TWAS), is committed to supporting scientific research and science policy by providing useful information on the present state of biotechnology and related issues. CoEBio in particular aims to contribute to the advancement of biotechnology development in developing countries, with the support of its biotechnology consultation system, its extensive training programs and through conducting multilateral R&D cooperation and tech-transfer programs.

"Biotechnology in Developing Countries: Growth and Competitiveness" is jointly issued by CoEBio and Clarivate Analytics (formerly the IP&Science business of Thomson Reuters). The report is designed to systematically analyze the status and productivity of biotechnology research in developing countries based on data obtained from scientific publications and patents documented in the 10 years from 2005-2014. To make the analysis thorough, the report divides the three main areas of biotechnology – medical, agricultural and industrial – into 32 specialized sub-disciplines.

Publication and patent data of 141 developing countries were retrieved from Clarivate Analytics (formerly the IP&Science business of Thomson Reuters) database and submitted for thorough comparative analyses. The status and competence of each developing country in each of the 32 subcategories of biotechnology can be easily drawn from the report.

This report is, to the best of my knowledge, the first extensive document summarizing the development status of a specific technology area in the developing world. It provides a strong,



valuable assessment of biotechnology activities in developing countries, as measured in scientific publications and patents. It may also serve as a macro reference for TWAS and UNESCO in understanding the current status and future trend of biotechnology development in developing countries. At the same time, through the analysis, it presents us the role and impact of international cooperation in the process.

I would like to congratulate CoEBio for successfully accomplishing this meaningful work. I hope that, in the near future, TWAS will be able to provide an overview of technology development in developing countries. Research work like this could also help governments and policymakers be better informed of technology development status in their country and, in turn, to incorporate this knowledge into sound policy-making.

"Biotechnology in Developing Countries: Growth and Competitiveness" could be valuable to governments and policymakers, related research and industry sectors, and relevant international bodies. It is our sincere hope the report will help to nurture a flourishing biotechnology sector in all developing countries and regions.

A handwritten signature in black ink, appearing to read 'C. Bai'.

BAI Chunli

President

Chinese Academy of Sciences (CAS)

The World Academy of Sciences for the advancement of science in developing countries (TWAS)

## Table of Contents

1	Executive Summary .....	5
2	Introduction.....	8
2.1	Overview .....	8
2.2	CAS-TWAS Centre of Excellence for Biotechnology .....	8
2.2.1	Chinese Academy of Sciences.....	8
2.2.2	Institute of Microbiology, Chinese Academy of Sciences.....	8
2.2.3	The World Academy of Sciences for the advancement of science in developing countries..	8
2.2.4	CAS-TWAS Centre of Excellence for Biotechnology .....	9
2.3	Clarivate Analytics.....	9
2.3.1	Clarivate Analytics Research Analytics .....	9
2.3.2	Clarivate Analytics Research Data Science & Evaluation team .....	9
2.3.3	Clarivate Analytics Intellectual Property Services .....	10
3	Data Sources, Indicators and Interpretation .....	11
3.1	Bibliometrics and citation analysis.....	11
3.2	Publication and Patent Data Source .....	12
3.2.1	Definition of Biotechnology Field .....	12
3.2.2	Country/territory Coverage.....	14
3.2.3	Data Collection.....	15
3.3	Methodology.....	16
3.3.1	Bibliometric indicators for publication analysis.....	16
3.3.2	Patent indicators .....	16
4	Overall trends of Biotechnology in Developing World .....	18
4.1	Research Publications, Overall and by Regional Grouping .....	18
4.2	Patent Activities, Overall and by Regional Grouping .....	24
4.3	Research Publications v.s. Patent Activities, Overall and by regional grouping.....	27
5	Sub-Categories of Biotechnology .....	28

5.1	Research Publications by Sub-Categories .....	28
5.2	Patent Activities by Sub-Categories.....	33
5.3	Research Publications v.s. Patent Activities by Sub-Categories.....	38
6	Overview of Biotechnology by Key Metrics .....	39
6.1	Research Publication Overview .....	39
6.2	Patent Activities Overview .....	46
7	Profile of Biotechnology by Developing Countries.....	53
7.1	Publication Profiles .....	53
7.2	Patent Profiles.....	182
	Annex 1 Countries/territories covered by TWAS regional offices.....	213
	Annex 2 Scientific Lagging Countries.....	217
	Annex 3 Keywords used in this project .....	218
	Annex 4 Bibliometrics and Citation analysis.....	225

## 1 Executive Summary

Biotechnology is the use of living systems and organisms to develop or make products, or "any technological application that uses biological systems, living organisms or derivatives thereof, to make or modify products or processes for specific use" (UN Convention on Biological Diversity, Art. 2<sup>1</sup>). Entwined with culture and socio-ethical values, biotechnology contributes to solving problems associated with population pressures, food insufficiency, nutritional deficiencies, and environmental degradation that impede national development in the developing world.

This report presents a quantitative analysis of research and development activities carried out by developing countries in Biotechnology field. Analyses presented in this report focus on research and patent publications between 2005 and 2014. The research and patent output are evaluated at both overall and regional level, and further disaggregated by 32 Biotechnology sub-categories (see Section 3.2.1), as well as all the developing countries defined by CAS-TWAS Centre of Excellence for Biotechnology (see Section 3.2.2).

### Summary of key findings

---

#### Research publications of developing countries in Biotechnology

- Research output of developing countries in the Biotechnology field has grown steadily from 2005 to 2014, with 117% increase in the number of papers (Section 4.1).
- Overall, the Biotechnology research outputs of developing countries are less cited than the world average, with normalized citation impact (0.83) remaining below 1.0 over the ten-year period. Furthermore, approximately eight to nine percent of papers are highly-cited, which is also slightly lower than the global baseline of 10% (Section 4.1).
- *East, South-East Asia and the Pacific Region* leads in the five regions and S&T Lagging Countries with the greatest output of papers and most significant increase in Biotechnology field. *Sub-Saharan Africa Region* and the *S&T Lagging Countries* benefit from higher international collaboration rate and resulted in fairly high research impact. In particular, *S&T Lagging Countries* has more than 85% of Biotechnology papers that are internationally collaborative (Section 4.1).
- Analyses at sub-category level show that the Biotechnology research by developing countries is concentrated in Medical, Clinical and Pharmaceutical Biotechnology fields. Of all 32 sub-categories, Infectious Diseases ranked the highest in terms of total output of papers (47,602), followed by Proteomics and Biomarkers (38,251)(Section 5.1).
- Plant Breeding, Biobased Fuels & Chemicals, Biosensors, Stem Cell Technology and Stem Cell Therapy are the five sub-categories with both citation impact reaching above the world average, and percentage of highly-cited papers higher than the global baseline of 10% (Section 5.1).
- Further breakdown by regions and countries/territories show that Proteomics and Biomarkers, Infectious Diseases, Enzymes, Chronic Diseases and Plant Transgenesis are the top

---

<sup>1</sup> <https://www.cbd.int/convention/articles/default.shtml?a=cbd-02>

subcategories in which all five regions and S&T Lagging Countries and top 10 countries/territories focused. In general, all regions and top 10 countries/territories share a similar interest across Biotechnology subcategories (Section 5.1).

- Analyses at country level show that China (78,263), India (24,081) and Brazil (17,769) lead in the output of Biotechnology research papers among all developing countries. All developing countries/territories increased their Biotechnology publications over the past ten years, except Venezuela. The countries<sup>2</sup> with highest publication growth, in terms of Compound Annual Growth Rate (CAGR)<sup>3</sup>, are Saudi Arabia, Iran, Pakistan, Malaysia and Algeria (Section 6.1 and Section 7.1).
- When considering the results and the individual bibliometric indicators presented for each country, the caveats made in Section 3.1, especially with regard to data sample size should always be borne in mind as the indicators can be skewed by low paper counts.

### Patent filings of developing countries in Biotechnology

- Total patent families of developing countries in Biotechnology field increased steadily from 2005 to 2014, with the average Compound Annual Growth Rate (CAGR) at 7.2% (Section 4.2).
- Similar to research publication activities, *East, South-East Asia and the Pacific Region* led in the five regions and S&T Lagging Countries, with the greatest output in the number of patent families. The regions with the lowest output of patents are *Arab Region*, *Sub-Saharan Africa Region*, and the *S&T Lagging Countries*. Among all five regions and S&T Lagging Countries, only *East, South-East Asia and the Pacific Region* shows a consistent growth (with 13% CAGR) between 2005 and 2013. (Section 4.2).
- Analyses at sub-category level show that the patent filings are most active in Industrial, Food and Environmental Biotechnology field. In particular, Enzymes (79,694) record the highest number of patent families among all 32 sub-categories, which takes more than 40% of the overall Biotechnology patent publications (Section 5.2).
- Microbial Fermentation Processes, Biorecycling and Bioproduct Production are the top areas with highest increase in terms of CAGR, indicating extensive technology/business opportunities underlying these sub-categories (Section 5.2).
- Further breakdown by regions and countries/territories show that Enzymes, Biobased Fuels & Chemicals, Infectious Diseases and Chronic Diseases are the top subcategories in which all regions and top 10 countries/territories share a similar focus. In particular, China has a higher than average focus on Bioproduct Production and Food Quality and Safety (Section 5.2).
- Analyses at country level show that China leads in all countries with 149,339 patent families, followed by India (15,420) and Mexico (14,574). Beside the overall output of patent, China is the only country which records a positive growth in CAGR (15%) and highest average patent

<sup>2</sup> Those have published at least 200 papers between 2005 and 2014.

<sup>3</sup> The Compound Annual Growth Rate was calculated by using  $\frac{\text{End year value}}{\text{Start year value}}^{\frac{1}{\# \text{ of years}}} - 1$ . The start year was the earliest year of publication, and the end year was the most recent publication year up to 2013. We have not considered the patent data for years 2014 for CAGR due to the lag in the publication of patents by the patent authorities.



remaining life of 15.41, indicating that China has made quite a significant achievement in Biotechnology R&D and new technologies have been constantly emerging in the ten-year period (Section 6.2 and Section 7.2).

- When considering the results and patent indicators presented for each country, data sample size should also be borne in mind for the same reason as stated before. Besides, the difference in lag between the filing of patent application and its publication by patent authorities is also one of the important reasons for the decline of patent activities in later years for most countries/territories<sup>4</sup>.

---

<sup>4</sup> The lag between the filing of patent application and its publication varies by patent authorities. For example in India, it may take several years for a patent application before it gets published, see <http://www.worldipreview.com/contributed-article/patent-delays-dealing-with-delays>. Same may hold true for other countries/territories, see a study ([http://www.scienpress.com/Upload/JAFB/Vol%203\\_5\\_9.pdf](http://www.scienpress.com/Upload/JAFB/Vol%203_5_9.pdf)) analysing the delay in patent publication in Taiwan(China) and its effect on growth.

## 2 Introduction

### 2.1 Overview

---

Chinese Academy of Sciences (CAS) (through its Institute of Microbiology, IMCAS) has commissioned a landscape and benchmarking analysis of research in Biotechnology for developing countries over the past 10 years. The aim of this analysis is to identify trends in Biotechnology research and development activities in developing countries and provide data which CAS can use to inform its strategic planning for the advancement of science in developing countries.

This commissioned project will:

- produce consistent analyses across research publications and patent data to provide a holistic picture of research & development in Biotechnology;
- analyse data over time to understand the trajectory of research performance and patent activity;
- evaluate Biotechnology as a whole and in sub-categories of this field;
- analyse trends in the developing world as a whole and in developing countries separately.

### 2.2 CAS-TWAS Centre of Excellence for Biotechnology

---

#### 2.2.1 Chinese Academy of Sciences

The Chinese Academy of Sciences (CAS) comprises a research and development network, a learned society and a system of higher education. It works with scientists and engineers from China and around the world to address both theoretical and applied problems using world-class scientific and management approaches. There were 124 Institutions directly under CAS at the end of 2012, with 104 research institutes, three universities and six Academic Divisions, 12 management organizations including the headquarters and branches, and three other units.

#### 2.2.2 Institute of Microbiology, Chinese Academy of Sciences

The Institute of Microbiology, Chinese Academy of Sciences (IMCAS) is the largest microbiological research institution in China. It was founded in 1958, through the merger of the Institute of Applied Mycology and the Beijing Laboratories of Microbiology, both of which were affiliated to the Chinese Academy of Sciences (CAS). IMCAS has become the nation's largest comprehensive research institution of microbiological science.

#### 2.2.3 The World Academy of Sciences for the advancement of science in developing countries

The World Academy of Sciences for the advancement of science in developing countries (TWAS) supports sustainable development and prosperity through research, education, policy and diplomacy. TWAS is a global scientific academy founded in 1983 by a distinguished group of scientists from the developing world, under the leadership of Abdus Salam, the Pakistani physicist and Nobel Prize winner. Today, TWAS has some 1,200 elected Fellows from more than 90 countries; 15 of them are Nobel laureates. The Academy is based in Trieste, Italy, on the campus of the Abdus Salam

International Centre for Theoretical Physics (ICTP). Since 1983, its mission has focused on supporting and promoting excellence in scientific research in the developing world and applying scientific and engineering research to address global challenges. TWAS receives core funding from the Government of Italy and essential programmatic funding from the Swedish International Development Cooperation Agency (Sida). The United Nations Educational, Scientific and Cultural Organization (UNESCO) administers TWAS funds and personnel. ([www.twas.org](http://www.twas.org))

#### **2.2.4 CAS-TWAS Centre of Excellence for Biotechnology**

The **CAS-TWAS Centre of Excellence for Biotechnology (CoEBio)** is one of 22 Centres of Excellence supported by the CAS and TWAS for the advancement of science in developing countries. Its main objective is to enhance the development of Biotechnology, to build scientific strength and drive innovation in developing countries. It does this by striving to improve and reinforce Biotechnology education, scientific communication, project cooperation and strategic intelligence analysis with the purpose of providing biotechnological solutions for addressing resources, energy, population, health and environment issues within developing countries.

### **2.3 Clarivate Analytics**

Clarivate Analytics, formerly the IP & Science business of Thomson Reuters, is the world's leading source of intelligent information for business and professionals. We combine industry expertise with innovative technology to deliver critical information to leading decision makers in the financial, legal, tax and accounting, healthcare, science and media markets, powered by the world's most trusted news organisation. Visit [WEBPAGE](#) for more information.

#### **2.3.1 Clarivate Analytics Research Analytics**

Clarivate Analytics Research Analytics is a suite of products, services and tools that provide comprehensive research analysis, evaluation and management. For over half a century we have pioneered the world of citation indexing and analysis, helping to connect scientific and scholarly thought around the world. Today, academic and research institutions, governments, not-for-profits, funding agencies, and all others with a stake in research need reliable, objective methods for managing and measuring performance. Visit [WEBPAGE](#) for more information.

#### **2.3.2 Clarivate Analytics Research Data Science & Evaluation team**

Clarivate Analytics Research Data & Evaluation team provides reporting and consultancy services within Research Analytics using customized analyses to bring together several indicators of research performance in such a way as to enable customers to rapidly make sense of and interpret a wide-range of data points to facilitate research strategy decision-making. We have extensive experience with databases on research inputs, activity and outputs and have developed innovative analytical approaches for benchmarking, interpreting and visualization of international, national and institutional research impact.

Our consultants have up to 20 years of experience in research performance analysis and interpretation. In addition, the Clarivate Analytics regional Sales team will provide effective project management and on-site support to maximize values of our projects and meet the expectations of the Chinese Academy of Sciences.

### 2.3.3 Clarivate Analytics Intellectual Property Services

Clarivate Analytics Global Intellectual Property (IP) Services has been providing professional IP services to clients for over 20 years, ranging from searching projects through to high-value consulting projects. Our teams provide robust reports to support thousands of clients making key business decisions throughout the world. Our services broadly support the IP lifecycle with a key focus on delivering high quality results that clients can rely on. Our footprint spans the world with major hubs in the US, UK, India and Japan. Clarivate Analytics' experience in the IP field began in the 1960s when Derwent Information started to create the value added database that is today's Clarivate Analytics Derwent World Patent Index®. Our patent analysis team has existed for over 10 years, serving over 120 organisations, many of whom are repeat clients. We produce landscape and other types of projects for diverse clients including global conglomerates, smash, academic institutions and government agencies. The group will conduct the intellectual property analysis.

## 3 Data Sources, Indicators and Interpretation

This Section outlines key concepts and methodology. A detailed description of the bibliometric methodology used by Clarivate Analytics is also provided in Annex 4.

### 3.1 Bibliometrics and citation analysis

---

Research evaluation is increasingly making wider use of bibliometric data and analyses. Bibliometrics is the analysis of data derived from publications and their citations. Publication of research outcomes is an integral part of the research process and is a universal activity. Consequently, bibliometric data have a currency across subjects, time and location that are found in few other sources of research-relevant data. The use of bibliometric analysis, allied to informed review by experts, increases the objectivity of, and confidence in, evaluation.

Research publications accumulate citation counts when they are referred to by more recent publications. Citations to prior work are a normal part of publication and reflect the value placed on a work by later researchers. Some papers get cited frequently and many remain uncited. Highly cited work is recognised as having a greater impact and Clarivate Analytics has shown that high citation rates are correlated with other qualitative evaluations of research performance, such as peer review<sup>5</sup>. This relationship holds across most science and technology areas and, to a limited extent, in social sciences and even in some humanities subjects.

Indicators derived from publication and citation data should always be used with caution. Some fields publish at faster rates than others and citation rates also vary. Citation counts must be carefully normalised to account for such variations by field. Because citation counts naturally grow over time, it is essential to account for growth by year. Normalisation is usually done by reference to the relevant global average for the field and for the year of publication.

Bibliometric indicators have been found to be more informative for core natural sciences, especially for basic science, than they are for applied and professional areas and for social sciences. In professional areas the range of publication modes used by leading researchers is likely to be diverse as they target a diverse, non-academic audience. In social sciences there is also a diversity of publication modes and citation rates are typically much lower than in natural sciences.

Bibliometrics analysis of patent data is derived from patent publications and their citations. There are marked similarities between literature bibliometrics and patent bibliometrics as both are applicable to the same set of problems. For a patent, the citation analysis provides information regarding reference of the patent by recent filings. The citation count of a patent family is helpful in estimating an impact that the invention is making. Again, citations to prior patent publication are a normal part of patent publications and reflect the value placed on a work by later inventors. Some patents get cited frequently and many may remain uncited. Highly cited patent publication is recognised as having a greater impact and Clarivate Analytics has shown that high citation rates are correlated with other qualitative evaluations of patent performance.

Bibliometrics work best with large data samples. As the data are disaggregated, so the relationship weakens. Average indicator values (e.g. of citation impact) for small numbers of publications can be

---

<sup>5</sup> Evidence Ltd. (2002) Maintaining Research Excellence and Volume: A report by Evidence Ltd to the Higher Education Funding Councils for England, Scotland and Wales and to Universities UK. (Adams J, et al.) 48pp.



skewed by single outlier values. At a finer scale, when analysing the specific outcome for individual departments, the statistical relationship is rarely a sufficient guide by itself. For this reason, bibliometrics are best used in support of, but not as a substitute for, expert decision processes. Well-founded analyses can enable conclusions to be reached more rapidly and with greater certainty, and are therefore an aid to management and used to instil increased confidence among stakeholders, but they cannot substitute for review by well-informed and experienced peers.

### 3.2 Publication and Patent Data Source

---

For this evaluation, bibliometric data are sourced from databases underlying the Clarivate Analytics Web of Science™, which gives access to conference proceedings, patents, websites, and chemical structures, compounds and reactions in addition to journals. It has a unified structure that integrates all data and search terms and therefore provides a level of comparability not found in other databases. It is widely acknowledged to be the world's leading source of citation and bibliometric data. The Clarivate Analytics Web of Science™ Core Collection is part of the Web of Science, and focuses on research published in journals and conferences in science, medicine, arts, humanities and social sciences. The authoritative, multidisciplinary content covers over 12,000 of the highest impact journals worldwide, including Open Access journals and over 150,000 conference proceedings. Coverage is both current and retrospective in the sciences, social sciences, arts and humanities, in some cases back to 1900. Within the research community, these data are often still referred to by the acronym 'ISI'. Clarivate Analytics has extensive experience with databases on research inputs, activity and outputs and has developed innovative analytical approaches for benchmarking and interpreting international, national and institutional research impact.

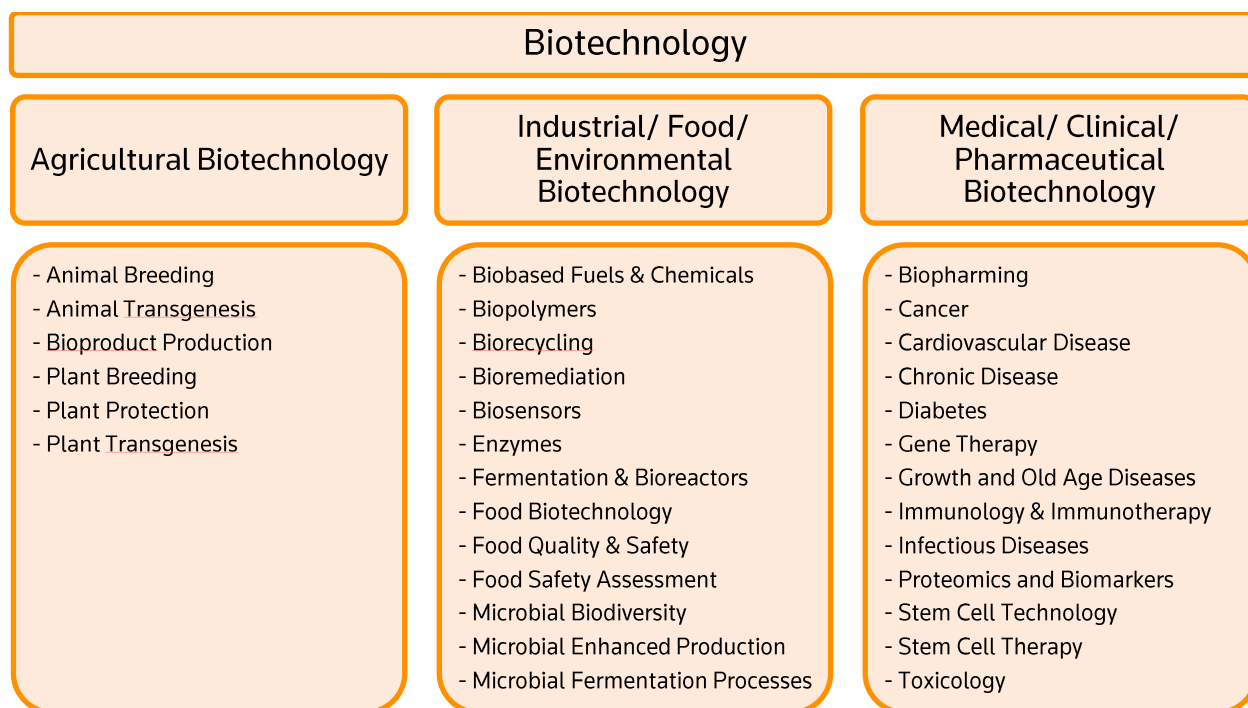
Patents data will be sourced from the Derwent World Patents Index and INPADOC-International Patent Documentation. Derwent World Patents Index is the world's most comprehensive database of enhanced patent documents. Subject experts from Clarivate Analytics correct, analyse, abstract and manually index every patent record, making it easier to quickly find the information needed to make informed decisions. Whether the patents are for their technical content, for business planning and development, for protecting the innovations in specific area or not, Derwent World Patents gives the most complete picture possible. INPADOC is an international patent collection produced and maintained by the European Patent Office (EPO). The INPADOC provides information about patent families, i.e., patent applications in different countries which claim the same priority and which normally disclose the same invention and it updated on a weekly basis.

#### 3.2.1 Definition of Biotechnology Field

A problem frequently encountered in the analysis of data about the research process is that of defining subject areas. For example, a funding body allocates money for chemistry but this goes to researchers in biology and engineering as well as to chemistry departments; clinicians publish in mathematics and education journals; publications in environmental journals come from a diversity of disciplines. This creates a problem when we try to define, for example, 'Biotechnology research'. Is this the work funded under Biotechnology programmes, the work of researchers in Biotechnology units or the work published in Biotechnology journals? Bibliometrics typically uses a predefined set of journals associated with a subject as a proxy for the body of research reflecting the field. This approach, however, is not always suitable; sometimes the subject matter for a particular analysis is not adequately defined by the available journal classifications.

For this evaluation, the publication data relating to Biotechnology will be collated using terms to search article titles, abstracts and keywords to identify relevant research publications – a methodology used by Clarivate Analytics in previous projects. This process is not an ‘exact science’ and relies upon interpretation and re-iteration to achieve results which give a dataset of ‘best fit’ which will adequately describe the research area whilst excluding publications of marginal relevance.

Article titles, abstracts and keywords will be searched using strings created from combinations of text words and text strings identified from these definitions to describe the sub-categories of Biotechnology. These sub-categories will be aggregated to define Biotechnology as a field in its entirety. The sub-categories to be analysed for this project have been arranged hierarchically, see Figure 3.2.1.



**Figure 3.2.1 Nested structure of the 32 Biotechnology sub-categories included in this project**

The keywords that we will use have been developed over the course of several projects and have been refined over several iterations. The same keyword sets will be used to identify relevant publications and patents (with small variations to allow for differences in indexing) to ensure that the analyses are consistent across these two content types. The list of keywords is provided in Annex 3 to this report.

### 3.2.2 Country/territory Coverage

This project will analyse, specifically:

- developing countries – defined as the 141 countries covered by TWAS; and
- developing countries by region – defined as the specific regions covered by the five TWAS regional offices, plus the 81 Science and Technology Lagging countries identified by TWAS.

Among the 141 developing countries/territories listed in Annex 1, publication data are available for 128 of the developing countries/territories; Macau and Hong Kong were combined with China Mainland together as a whole China, and the following 11 countries haven't published any papers in Biotechnology field.

**Table 3.2.2 List of countries/territories not included in publications analyses due to lack of presence in Biotechnology field**

Country
Aruba
Guinea Equatorial/Equatorial Guinea
Falkland Islands
Saint Vincent and the Grenadines
Palestinian Aut. Terr.
Tuvalu
Timor-Leste
Kiribati
Tonga
Samoa
Republic of South Sudan

Patent data are available for the following 30 developing countries/territories:

**Table 3.2.2. List of countries/territories included in patent analyses**

Country	Country	Country
Algeria	Egypt	Panama
Argentina	El Salvador	Peru
Brazil	Guatemala	Philippines
Chile	Honduras	South Africa
China	India	Taiwan, China
Colombia	Indonesia	Tajikistan
Costa Rica	Malaysia	Thailand
Cuba	Mexico	Turkey
Dominican Republic	Morocco	Uruguay
Ecuador	Nicaragua	Vietnam

### 3.2.3 Data Collection

In this report, all the publication data was collected in July 2015, by using strings created from combinations of text words and text strings (see Annex 3) to search publication titles, abstracts and author keywords and keywords plus indexed in SCI-EXPANDED, SSCI and A&HCI of Web of Science™ Core Collection.

More specifically, only the publications with document types of articles and reviews are included in the analyses, and editorials, meeting abstracts or other types of publications are not included. More details on methodology can be found in the following Section 3.3.

Furthermore, all the patent data was collected in July 2015, by using strings created from combinations of text word, text strings and relevant patent classifications (see Annex 3) to search patent titles, abstracts and technical specifications of the patent filings for the 30 countries/territories which are part of the study.

### 3.3 Methodology

---

#### 3.3.1 Bibliometric indicators for publication analysis

**Papers/publications:** Clarivate Analytics abstract publications including editorials, meeting abstracts and book reviews as well as research journal articles. The terms ‘paper’ and ‘publication’ are often used interchangeably to refer to printed and electronic outputs of many types. In this document the term ‘paper’ has been used exclusively to refer to substantive journal articles, reviews and some proceedings papers and excludes editorials, meeting abstracts or other types of publication. **Papers** are the subset of publications for which citation data are most informative and which are used in calculations of citation impact. Furthermore, papers tend to be those that report research findings and they are peer-reviewed, whereas other publication types often cover other materials and are not reviewed in the same way, which also explains why they are most often used in research evaluations as an internationally accepted best practice.

**Citations:** The citation count is the number of times that a citation has been recorded for a given publication since it was published. Not all citations are necessarily recorded since not all publications are indexed. The material indexed by Clarivate Analytics, however, is estimated to attract about 95% of global citations.

**Citation impact:** ‘Citations per paper’ is an index of academic or research impact (as compared with economic or social impact). It is calculated by dividing the sum of citations by the total number of papers in any given dataset (so, for a single paper, raw impact is the same as its citation count). Impact can be calculated for papers within a specific research field such as Clinical Neurology, or for a specific institution or group of institutions, or a specific country. Citation count declines in the most recent years of any time period as papers have had less time to accumulate citations (papers published in 2007 will typically have more citations than papers published in 2010).

**Field-normalised citation impact (NCI<sub>F</sub>):** Citation rates vary between research fields and with time, consequently, analyses must take both field and year into account. In addition, the type of publication will influence the citation count. For this reason, only citation counts of papers (as defined above) are used in calculations of citation impact. The standard normalisation factor is the world average citations per paper for the year and journal category in which the paper was published. This normalisation is also referred to as ‘rebasing’ the citation count.

**Mean normalised citation impact (mNCI):** The mNCI indicator for any specific dataset is calculated as the mean of the NCI<sub>F</sub> of all papers within that dataset.

**Percentage of highly-cited papers:** For the purpose of this report, highly-cited papers have been defined as those articles and reviews which belong to the world’s top 10% of papers in that journal category and year of publication, when ranked by number of citations received. A percentage that is above 10 indicates above-average performance.

#### 3.3.2 Patent indicators

**Number of Patent Families:** Total Patent Families are INPADOC patent families. An INPADOC patent family is defined as comprising all the documents sharing directly or indirectly (e.g. via a third document) at least one priority. This includes all the patent documents resulting from a patent



application submitted as a first filing with a patent office and from the same patent application filed within the priority year with a patent office in any other country.<sup>6</sup>

**Geographic Filing breadth:** measures the number of different geographic locations or legal jurisdictions into which the patent has been filed. This measurement is strongly associated with increasing costs and, therefore, a method of identifying patented subject matter that is more strategic (or conversely, more speculative) to the patent applicant.

**Grant success/grant commitment/granted patent locations:** measures and scores the patent family based on the location and number of granted patents that exist within the family, for example granted in the United States, at the European Patent Office, in Japan. Some locations are more important than others, particularly when assessing enforcement potential, licensing opportunity or market applicability, and this can be incorporated into the model. For the purpose of this study, the grant success is also calculated for the patent families of the individual countries which were analysed in this study.

**Frequency of citation by downstream patent applications:** based on the INPADOC, this measurement acts as proxy for the impact a particular invention or patent family is having within its technical field. The metric necessarily uses frequency rather than the raw number of citation events to take account of age bias. On its own, citation is not necessarily an indicator of intrinsic value of the IP but, in combination with other measurements, can identify potentially high value IP. Furthermore, where individual patent families or portfolios are highly cited by third parties but lacking in other metrics, it can point to IP strategy changes or opportunities to exploit the IP strategies of others.

**Technical Breadth:** measures the number of different technology areas in which a patent family or invention is relevant. This is based on the DWPI<sup>SM</sup> Classification, a manually applied classification, therefore broader and more insightful than other classification schemes. All other items being equal, a technically broad patent provides the owner with greater assertion opportunity, and therefore maps to more valuable IP.

**Age of the patent:** based on the amount of remaining enforcement opportunity the patent has left.

---

<sup>6</sup> [http://worldwide.espacenet.com/help?topic=patentfamily&locale=en\\_EP&method=handleHelpTopic](http://worldwide.espacenet.com/help?topic=patentfamily&locale=en_EP&method=handleHelpTopic)

## 4 Overall trends of Biotechnology in Developing World

This section analyses the trends in research publication and patent activity between 2005 and 2014. The publications and patents are aggregated for all Biotechnology research across all developing countries and by regional groupings. The regional groupings are:

- Latin America and the Caribbean Region
- East, South-East Asia and the Pacific Region
- Arab Region
- Central and South Asia Region
- Sub-Saharan Africa Region
- S&T Lagging Countries

The countries which make up these region groupings are listed in Annex 1 and Annex 2.

### 4.1 Research Publications, Overall and by Regional Grouping

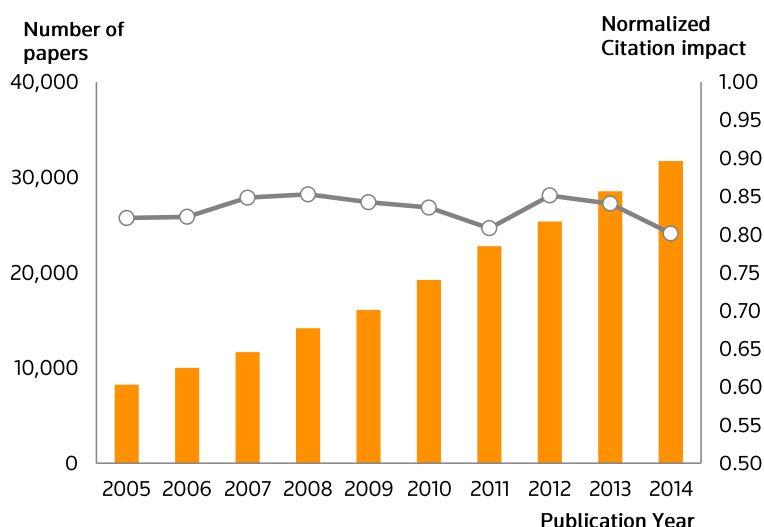


Figure 4.1.1 Number and normalized citation impact of Biotechnology papers, 2005-2014

The number of papers in Biotechnology field has increased steadily every year between 2005 and 2014. Overall, there has been a 117% increase in the number of papers published in this field. The normalized citation impact of these papers was between 0.80 and 0.85 during this period, meaning they're less cited than the world average. In general, this indicates that as the number of papers has grown the impact of these papers has remained the same.

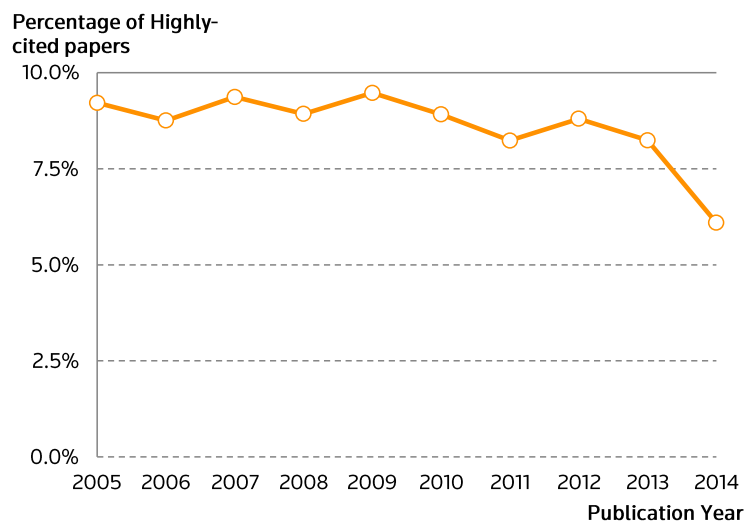


Figure 4.1.2 Percentage of highly-cited Biotechnology papers, 2005-2014

The percentage of highly-cited papers has remained steady around 9% between 2005 and 2013, a slightly lower than world average level of 10%. Since citations take time to accrue, it is not surprising that the percentage of highly-cited papers for the two most recent years of analysis, 2013 and 2014 are lower.

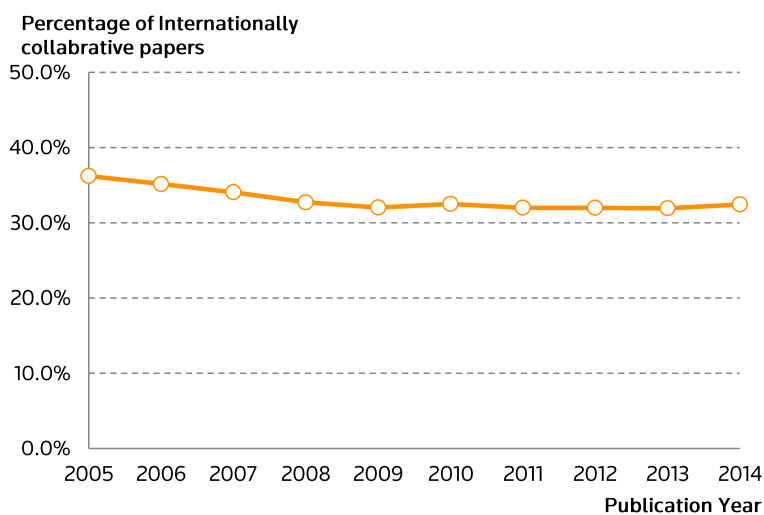


Figure 4.1.3 Percentage of internationally collaborative Biotechnology papers, 2005-2014

A paper is internationally collaborative if there are multiple author affiliations from different countries for that paper. The percentage of internationally collaborative papers in Biotechnology has essentially remained the same around 30% for all developing countries, between 2005 and 2014. This indicates that about one out of three papers are internationally collaborative in this field.

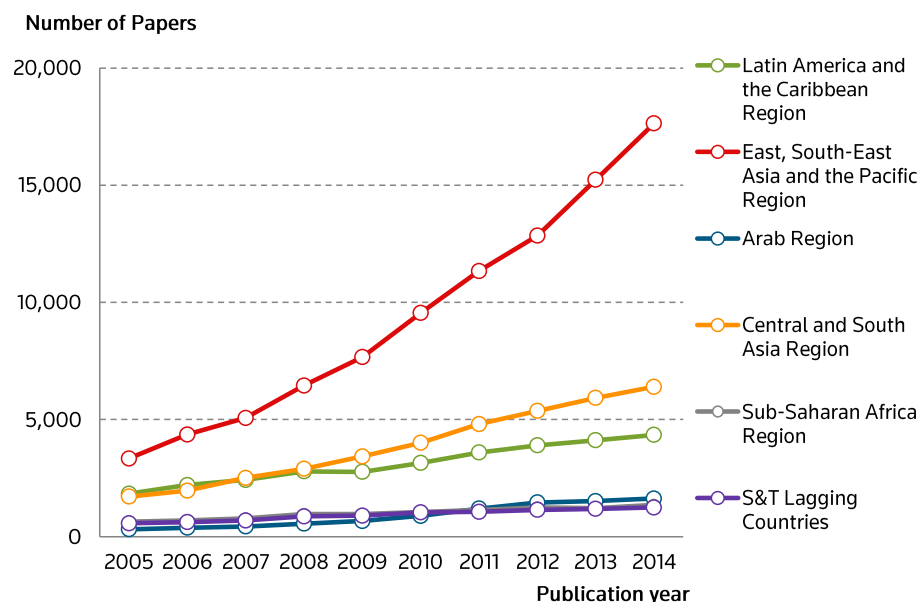


Figure 4.1.4 Number of papers from six regional areas in Biotechnology, 2005-2014

The *East, South-East Asia and the Pacific Region* had the highest number of papers and increased rapidly from 2005 to 2014. The regions with the lowest number of papers were *Arab Region*, *Sub-Saharan Africa Region*, and the *S&T Lagging Countries*.

The *Central and South Asia Region* was behind *Latin America and Caribbean Region* from 2005 to 2008, but has since then taken over, and between 2009 and 2014 was second only to *East, South-East Asia and the Pacific Region* in the overall output of Biotechnology papers.

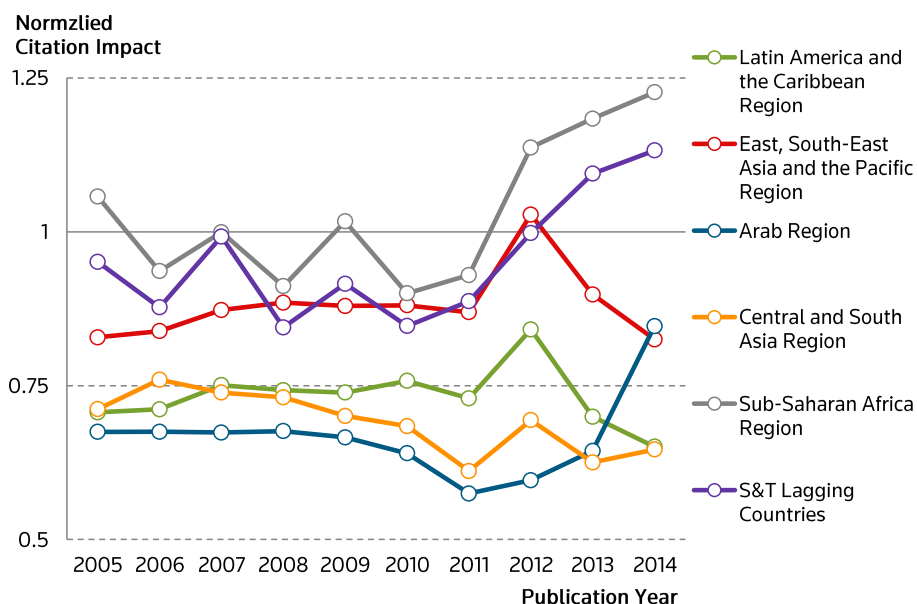


Figure 4.1.5 Normalized Citation impact from six regional areas in Biotechnology, 2005-2014

The *Sub-Saharan Africa Region* had the highest normalized citation impact of all five regions and S&T Lagging Countries between 2005 and 2014. The *S&T Lagging Countries* ranked the second highest in the same indicator for many years. This can be understood as follow: both of the regions have benefitted from the higher than average percentage of international collaboration (see Figure 4.1.3 and Figure 4.1.7). In fact, the average normalized citation impact of the non-collaborative papers from the *S&T Lagging Countries* is only 0.35.

The *East, South-East Asia and the Pacific Region* had the third highest normalized citation impact between 2005 and 2014, but were also the regions with the highest output of Biotechnology papers (see Figure 4.1.1). Analysing these two figures together indicates that the *East, South-East Asia and the Pacific Region* is the regional leader of Biotechnology research in developing world.



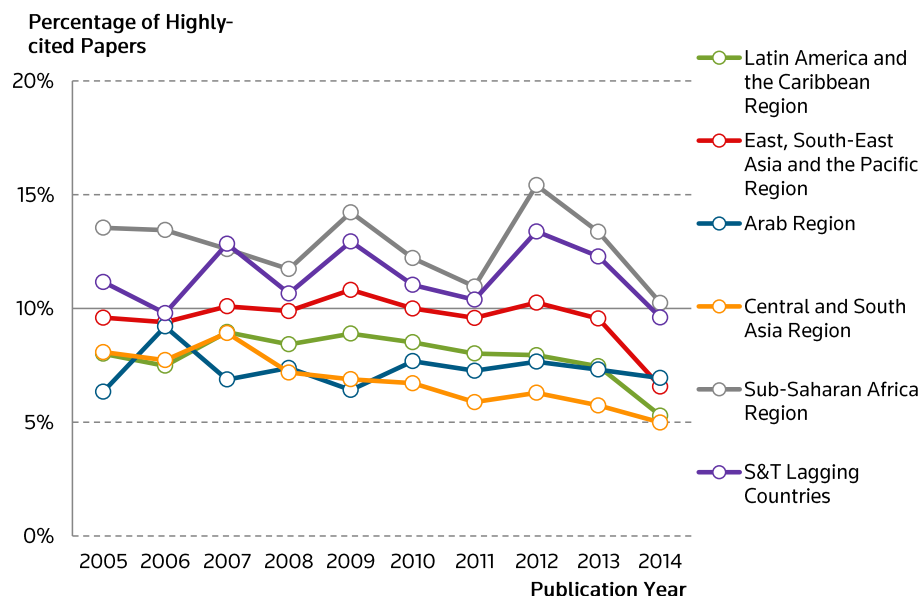


Figure 4.1.6 Percentage of highly-cited papers from six regional areas in Biotechnology, 2005-2014

The *Sub-Saharan Africa Region* and the *S&T Lagging Countries* had the highest percentages of highly-cited research papers. As shown in Figure 4.1.5, this is also mostly due to the higher international collaboration rate. Indeed, the percentage of highly-cited papers of the *S&T Lagging Countries* from non-internationally collaborative papers is 1.28% for their research published between 2005 and 2014.

The *East, South-East Asia and the Pacific Region* had the third highest percentage of highly-cited papers between 2005 and 2013, but was also the region with the highest number of publications (see Figure 4.1.4). Given the large volume of research output, *East, South-East Asia and the Pacific Region* performs fairly well in terms of highly-cited research in the field of Biotechnology.

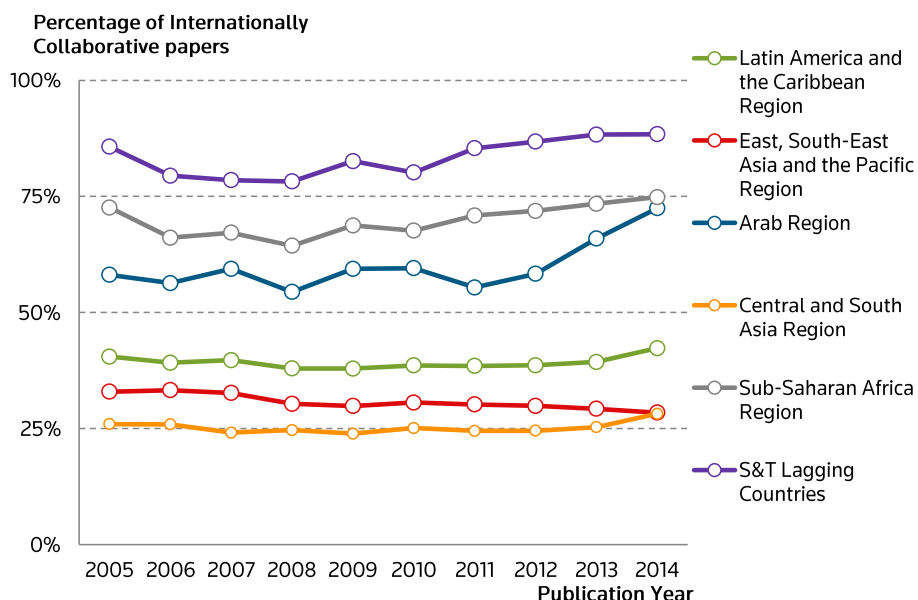


Figure 4.1.7 Percentage of internationally collaborative papers from six regional areas in Biotechnology, 2005-2014

International research collaboration is a rapidly growing element of research activity. The reasons for this include increasing access to facilities and resources, as well as access to knowledge, people and expertise.

It can be seen from Figure 4.1.7 that, the *S&T Lagging Countries* had the highest percentage of internationally collaborative papers (90%). The second highest was from *Sub-Saharan Africa Region*. While *East, South-East Asia and the Pacific Region* and *Central and South Asia Region* had the highest number of papers of all five regions and *S&T Lagging Countries*, they ranked fifth and sixth respectively in terms of international collaboration rate in Biotechnology.

This analysis highlights again that international collaboration is associated with an increase in the number of citations received by the research papers. And for both the *East, South-East Asia and the Pacific Region* and *Central and South Asia Region*, strengthening the international collaboration might be a good solution to increase the impact of their research, as measured by normalized citation impact and percentage of highly-cited papers.

## 4.2 Patent Activities, Overall and by Regional Grouping

The following graph shows the filing trend of patents based on the priority date of the patent families from 2005 to 2014.

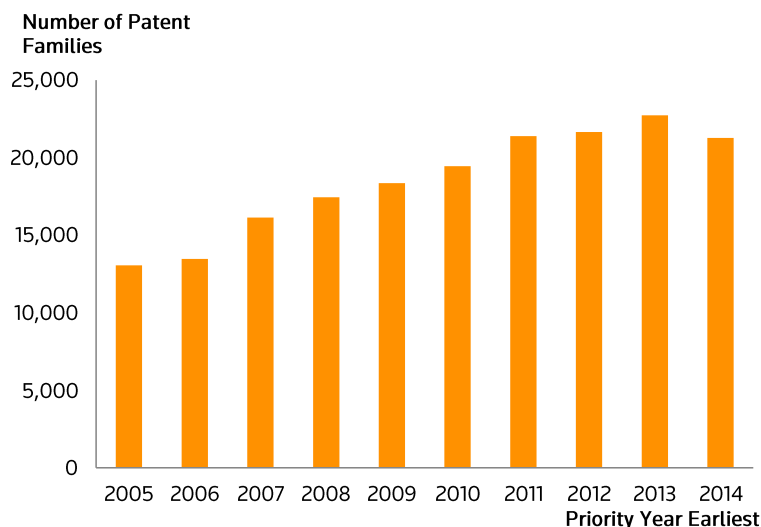


Figure 4.2.1 Number of patent families in overall Biotechnology field, 2005 - 2014

The overall number of patent filings in Biotechnology in developing countries has grown from 13,058 in 2005 to 21,263 in 2014. The graph shows a slight dip in 2014, which may be attributed to the 18 month gap between the filing and publication of the patent publications by the patent authorities. Overall, the Compound Annual Growth Rate (CAGR) of patent filings between 2005 and 2013 reached 7.2%.

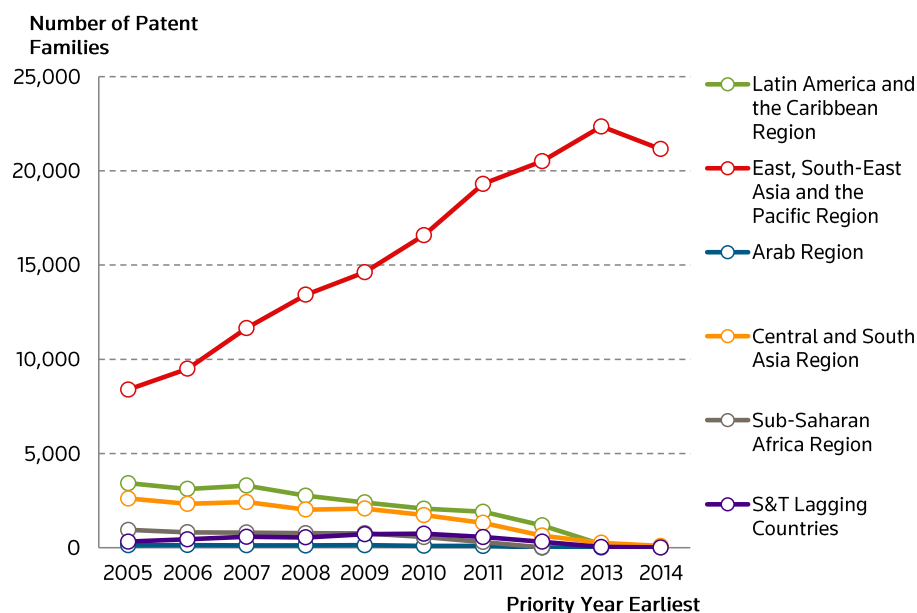


Figure 4.2.2 Number of patent families from six regional areas in Biotechnology, 2005-2014

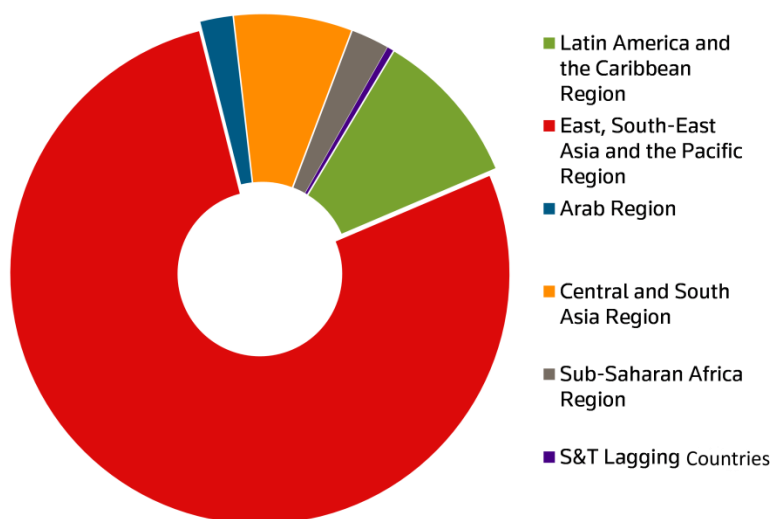


Figure 4.2.3 Number of patent families from six regional areas in Biotechnology, 2005-2014 in aggregate

As shown in Figure 4.2.2 and Figure 4.2.3, among the regional groupings, *East, South-East Asia and the Pacific Region* had the highest number of patent families from 2005 to 2014. The region shows a slight dip in 2014, which may be due to the 18-month time gap between the filing and publication of the patent. The regions with the lowest number of patent filings were *Arab Region*, *Sub-Saharan Africa Region* and the *S&T Lagging Countries*.

The figure below represents the Compound Annual Growth Rate (CAGR) of patent filings among regional groupings between 2005 and 2013. The CAGR is calculated to show the growth rate over the years in the filing of patents. However, we have not considered the data for years 2014 for CAGR due to the lag in the publication of patents by the patent authorities as mentioned before.

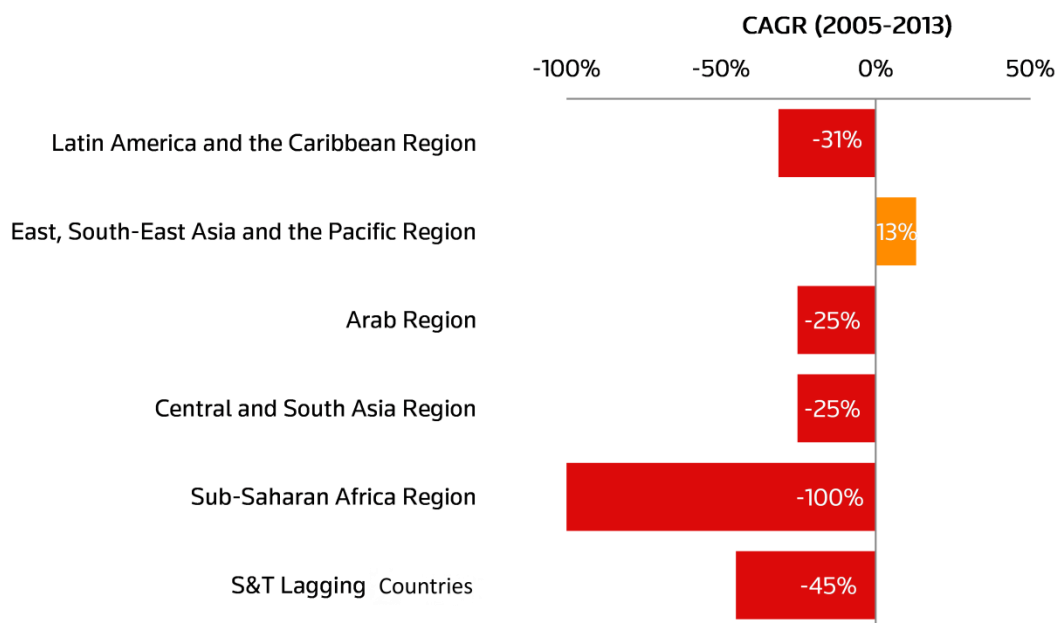


Figure 4.2.4 Compound annual growth rate of patent filings in overall Biotechnology field between 2005 and 2013

As can be seen in Figure 4.2.4, *East, South-East Asia and the Pacific Region* was the only region to record a positive growth. In comparison, the other the regional offices show a decline in the filing trend over the years. The *Sub-Saharan Africa Region* shows the highest decline in growth.

### 4.3 Research Publications v.s. Patent Activities, Overall and by regional grouping

---

By comparing the major findings from the overall analyses on research publications and patent activities, it can be seen that:

- The research output and patent filings of developing countries in Biotechnology field increased steadily from 2005 to 2014.
- *East, South-East Asia and the Pacific Region* led in the six regional groupings, and has the greatest output in both number of papers and patent families.
- All five regions and S&T Lagging Countries had increasing output of papers in Biotechnology, while only *East, South-East Asia and the Pacific Region* had the positive growth rate in patent filings.
- With the lowest publication output, *Sub-Saharan Africa Region* and the *S&T Lagging Countries* benefitted from higher international collaboration rate and had fairly high-level of research impact.
- *Central and South Asia Region* had more Biotechnology research publications than *Latin America and the Caribbean Region* over the last ten years, while the latter filed more patents in Biotechnology.

## 5 Sub-Categories of Biotechnology

This section presents the analyses of Biotechnology publications and patents for the 10-year time frame, disaggregated by three sub-fields and 32 sub-categories.

The analyses on research publications of 128 developing countries/territories focused on the output of papers, normalized citation impact, percentage of highly-cited papers, percentage of internationally collaborative papers, and normalized citation impact of internationally collaborative papers, for each sub-category. At the same time, the analyses of patent activities in 30 countries/territories focused on the patent filings, compound annual growth rate, patent remaining life, time-weighted patent citation, grant success and geographical coverage of patent applications.

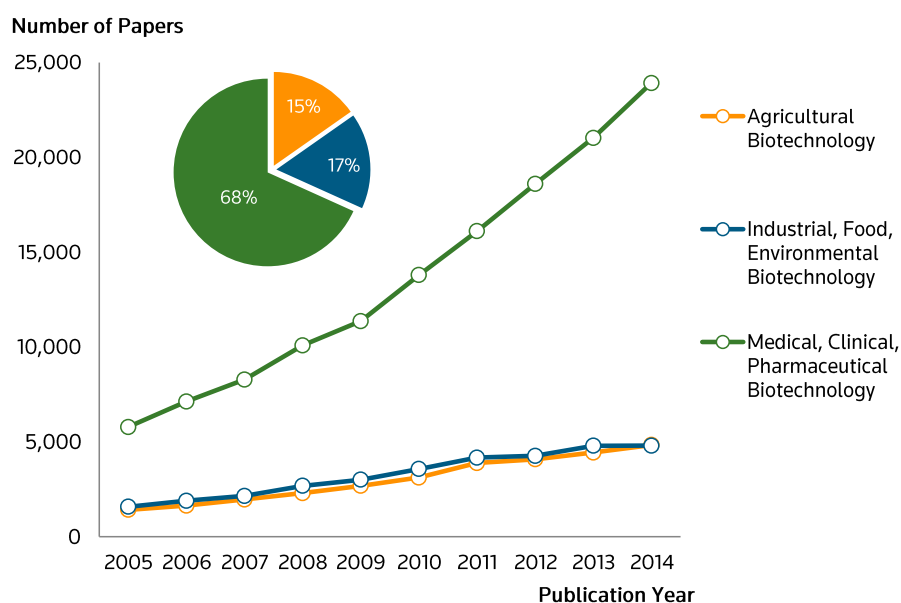
### 5.1 Research Publications by Sub-Categories

Table 5.1.1 shows the summary bibliometric indicators for overall Biotechnology publications affiliated to 128 developing countries/territories.

**Table 5.1.1 Bibliometric analysis of the overall Biotechnology field, 2005-2014**

Biotechnology	Number of papers	Average NCI	Highly-cited papers (%)	Internationally Collaborative Papers (%)	Average NCI of internationally collaborative papers
Overall	187,692	0.83	8.3%	32.6%	1.15

In 2005-2014, 187,692 Biotechnology papers were published by developing countries in the past ten years. The normalized citation impact was 0.83, indicating these papers were cited less the world average. There are 8.3% of papers which could be ranked as the world's top 10% of papers in that journal category and year of publication. Overall, nearly one out of three papers was internationally collaborative. International collaboration has been shown to be associated with an increase in citation impact, as these internationally collaborative papers were cited more than the world average.



**Figure 5.1.1 Trends in Biotechnology publications by three sub-fields, 2005-2014**



The following Table 5.1.2 shows the summary bibliometric indicators for each of the 32 sub-categories in Biotechnology.

Table 5.1.2 Summary bibliometric indicators of publications by 32 Biotechnology sub-categories, 2005-2014<sup>7</sup>

Biotechnology subcategories		Number of papers	Average NCI	Highly-cited papers (%)	Internationally Collaborative Papers (%)	Average NCI of internationally collaborative papers
Agricultural	Animal Breeding	1,272	0.81	9.0	40.9	1.24
	Animal Transgenesis	4,861	0.71	5.8	32.3	1.04
	Bioproduct Production	1,467	0.78	7.6	27.1	1.06
	Plant Breeding	8,887	1.01	10.7	29.6	1.36
	Plant Protection	5,731	0.92	8.3	33.8	1.26
	Plant Transgenesis	11,344	0.96	9.9	34.9	1.28
Industrial, Food, Environmental	Biobased Fuels & Chemicals	4,435	1.26	15.7	25.5	1.44
	Biopolymers	3,043	0.85	7.9	25.0	1.10
	Biorecycling	138	0.73	5.1	23.2	0.98
	Bioremediation	3,357	0.81	7.0	24.5	1.01
	Biosensors	576	1.16	15.3	27.6	1.28
	Enzymes	13,387	0.70	6.2	25.9	0.93
	Fermentation and Bioreactors	159	0.67	4.4	28.9	1.15
	Food Biotechnology	4,004	0.68	6.4	30.5	0.84
	Food Quality and Safety	1,161	0.70	8.2	33.7	1.02
	Food Safety Assessment	734	0.78	8.5	31.7	1.10
	Microbial Biodiversity	2,475	0.91	9.1	34.9	1.15
	Microbial Enhanced Production	1,910	0.89	9.3	25.6	1.18
	Microbial Fermentation Processes	1,326	0.83	8.0	31.0	1.07
	Biopharming	259	0.81	8.1	31.7	1.12
Medical, Clinical, Pharmaceutical	Cancer	11,774	0.93	10.1	33.1	1.23
	Cardiovascular Diseases	5,567	0.62	5.4	27.4	0.92
	Chronic Diseases	15,410	0.76	7.5	28.1	1.10
	Diabetes	8,883	0.75	7.4	26.7	1.07
	Gene Therapy	4,042	0.81	7.1	30.6	1.13
	Growth and Old Age Diseases	6,168	0.76	7.9	31.5	1.11
	Immunology and Immunotherapy	9,706	0.77	6.9	29.9	1.22
	Infectious Diseases	47,602	0.81	8.5	39.5	1.14
	Proteomics and Biomarkers	38,251	0.84	8.1	32.7	1.15
	Stem Cell Technology	9,649	1.02	12.0	34.3	1.45
	Stem Cell Therapy	614	1.00	11.4	35.2	1.36
	Toxicology	1,027	0.86	8.1	30.1	1.26

<sup>7</sup> The cells are highlighted to reflect the distinction between the listed indicators across 32 subcategories. Overlaps between subcategories are to be expected due to the definition and methodology of categorization mentioned in Section 3.2. This also holds true for the following Table 5.1.3 and Table 5.1.4.

It can be seen from Figure 5.1.1 that the Medical, Clinical, Pharmaceutical Biotechnology accounted for over two-thirds of overall research publications, and presented an exponential growth in the last ten years.

Of all 32 sub-categories, Infectious Diseases (47,602) ranked the highest in terms of total output of papers, followed by Proteomics and Biomarkers (38,251). Plant Transgenesis (11,344) accounted for 34% of the total publication output in Agricultural Biotechnology field, which also ranked highest in this sub-field. Enzymes is the category which accounted over one third of total publications in Industrial, Food, and Environmental Biotechnology field.

5 out of 32 sub-categories had normalized citation impact higher than 1, indicating that the biotechnology research of developing countries in these sub-categories were cited more than the world average.

The percentage of highly-cited papers was closely tied to normalized citation impact, where the same five sub-categories had higher than 10% of highly-cited papers; and they are Plant Breeding, Biobased Fuels & Chemicals, Biosensors, Stem Cell Technology and Stem Cell Therapy.

Overall, nearly one third of biotechnology publications are internationally collaborative. Animal Breeding had the highest international collaboration rate (40.9%), followed by Infectious Diseases (39.5%). Biorecycling ranked lowest in both international collaboration rate (23.2%) and overall output (138).

The internationally collaborative papers were cited more than the world average for all sub-categories except for Food Biotechnology, Cardiovascular Diseases, Enzymes and Biorecycling.

Table 5.1.3 Total papers by 32 Biotechnology sub-categories for each region, 2005-2014<sup>8</sup>

Biotechnology subcategories		Latin America and the Caribbean Region	East, South-East Asia and the Pacific Region	Arab Region	Central and South Asia Region	Sub-Saharan Africa Region	S&T Lagging Region
Agricultural	Animal Breeding	436	512	63	237	137	140
	Animal Transgenesis	558	3,280	161	910	80	85
	Bioproduct Production	233	801	84	290	144	126
	Plant Breeding	1,229	4,144	832	3,036	309	382
	Plant Protection	1,399	2,193	327	1,707	580	554
	Plant Transgenesis	2,417	5,016	567	3,254	944	898
Industrial/Food/Environmental	Biobased Fuels & Chemicals	886	2,489	141	986	163	155
	Biopolymers	595	1,662	118	696	78	64
	Biorecycling	15	67	10	44	6	6
	Bioremediation	610	1,383	263	1,082	206	165
	Biosensors	70	318	41	142	27	21
	Enzymes	2,772	6,628	762	3,944	480	419
	Fermentation and Bioreactors	30	75	9	49	6	7
	Food Biotechnology	731	2,483	133	679	160	150
	Food Quality and Safety	289	355	136	354	129	140
	Food Safety Assessment	218	289	54	143	80	93
	Microbial Biodiversity	572	1,225	148	552	92	68
	Microbial Enhanced Production	371	747	242	548	122	117
Medical/Clinical/Pharmaceutical	Microbial Fermentation Processes	260	652	78	265	134	100
	Biopharming	78	90	12	63	30	15
	Cancer	1,323	7,748	715	2,188	178	145
	Cardiovascular Diseases	1,234	2,898	370	1,082	170	131
	Chronic Diseases	2,601	8,804	1,096	2,891	582	591
	Diabetes	1,482	4,247	830	2,449	199	191
	Gene Therapy	398	3,164	105	390	44	27
	Growth and Old Age Diseases	1,122	3,424	332	1,368	137	114
	Immunology and Immunotherapy	2,004	5,005	664	2,027	428	349
	Infectious Diseases	10,449	20,681	2,638	10,648	6,981	6,388
	Proteomics and Biomarkers	5,815	24,012	1,480	7,288	965	873
	Stem Cell Technology	814	7,335	335	1,299	84	63
	Stem Cell Therapy	83	398	24	125	2	5
	Toxicology	249	407	92	256	71	68

<sup>8</sup> The cells are highlighted to reflect the relative research focus in each region.

Table 5.1.4 Total papers by 32 Biotechnology sub-categories for top 10 countries/territories, 2005-2014<sup>9</sup>

Biotechnology subcategories		Argentina	Brazil	China	India	Iran	Mexico	South Africa	Taiwan, China	Thailand	Turkey
Agricultural	Animal Breeding	37	243	397	115	55	44	60	33	23	33
	Animal Transgenesis	61	293	2,768	506	199	118	42	329	89	113
	Bioproduct Production	18	139	669	162	56	29	50	66	22	27
	Plant Breeding	208	529	3,710	1,355	484	277	113	145	81	334
	Plant Protection	128	768	1,782	1,031	172	230	184	132	94	162
	Plant Transgenesis	403	1,063	4,078	1,882	321	528	369	378	164	355
Industrial/Food/Environmental	Biobased Fuels & Chemicals	67	547	1,788	662	122	122	96	262	157	146
	Biopolymers	118	295	1,259	489	66	87	43	226	79	89
	Biorecycling	1	8	49	37	2	3	2	9	5	3
	Bioremediation	96	253	1,051	778	105	128	80	116	61	77
	Biosensors	21	27	265	66	33	5	15	21	11	38
	Enzymes	399	1,571	5,078	2,831	398	406	219	779	325	475
	Fermentation and Bioreactors	9	11	46	20	11	10	4	14	4	12
	Food Biotechnology	90	452	2,114	380	83	90	73	186	68	102
	Food Quality and Safety	53	151	231	173	68	52	30	30	34	74
	Food Safety Assessment	42	101	221	68	23	41	20	23	15	32
	Microbial Biodiversity	71	265	1,019	375	52	102	54	78	58	87
	Microbial Enhanced Production	48	214	533	313	122	69	36	62	63	74
	Microbial Fermentation Processes	49	131	496	168	26	39	61	68	41	50
Medical/Clinical/Pharmaceutical	Biopharming	22	20	78	46	8	23	19	6	1	4
	Cancer	217	698	6,281	1,263	350	204	101	1,128	129	472
	Cardiovascular Diseases	150	789	2,352	547	131	131	97	386	51	335
	Chronic Diseases	262	1,554	7,072	1,412	419	340	185	1,217	219	863
	Diabetes	175	866	3,424	1,352	343	237	83	544	99	624
	Gene Therapy	89	182	2,779	183	101	76	30	312	35	96
	Growth and Old Age Diseases	164	591	2,913	814	208	148	86	367	56	275
	Immunology and Immunotherapy	295	937	4,084	777	474	317	208	593	160	694
	Infectious Diseases	1,578	5,230	14,252	6,296	1,491	1,390	2,436	3,004	1,624	1,835
	Proteomics and Biomarkers	909	2,848	19,461	4,789	860	924	505	3,211	708	1,144
	Stem Cell Technology	90	485	6,320	637	418	119	61	775	108	201
	Stem Cell Therapy	4	58	331	56	41	5	2	48	4	20
	Toxicology	49	116	294	148	39	32	24	64	19	46

<sup>9</sup> The cells are highlighted to reflect the relative research focus in each country/territory.

## 5.2 Patent Activities by Sub-Categories

The table below represents all relevant patent indicators<sup>10</sup> for the overall Biotechnology field. Again, the CAGR was calculated between 2005 and 2013 by taking the lag in patent publications into consideration.

Table 5.2.1 Patent activity of overall Biotechnology field (2005-2014)

Biotechnology	Total Patent Families	CAGR % (2005-2013)	Avg. Patent Remaining Life	Avg. Time-weighted Patent Citation	Grant Success Rate %	Avg. Geo-coverage of Patent Application
Overall	184,957	7.2	14.94	0.47	41.1	2

The total number of families recorded for the overall Biotechnology field during the period of the study is 184,957, with a CAGR of 7.2%. It shows that the overall Biotechnology field has shown a constant growth over the years. The average remaining life of patents comes about to 15 years, which illustrated that most of the patent families have a substantial lifetime till expiration. The average time-weighted citation was recorded to be 0.47. The grant success rate for the overall biotechnology sector recorded is 41.1%, again it shows that the grant success is towards a lower margin, however, keeping the lag period between the filing and grant of a patent in consideration, this rate may increase, once the filings in recent years receive a grant.

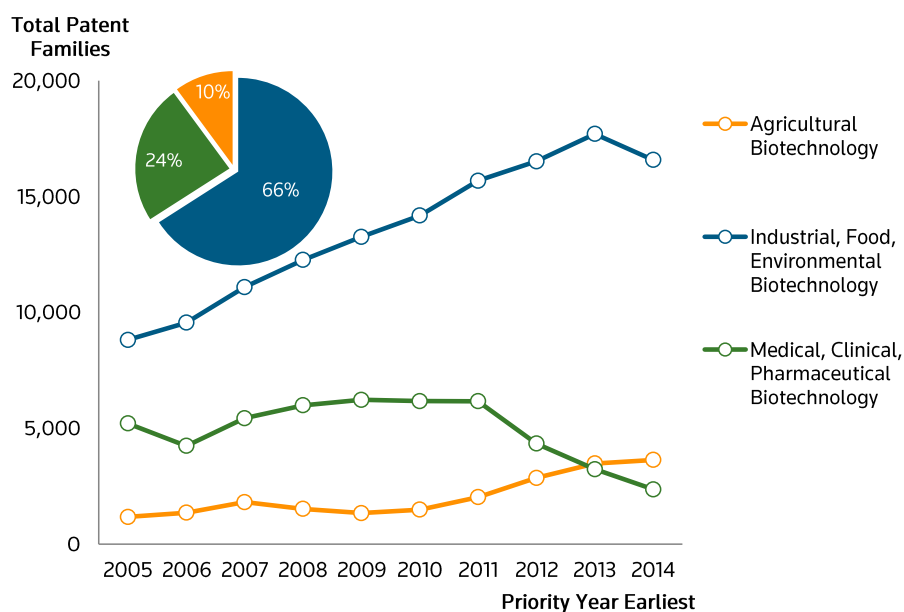


Figure 5.2.1 Trends in Biotechnology patent activities by three sub-fields, 2005-2014

Table 5.2.2 shows the number of patent families recorded for each of 32 sub-category along with various parameters in aggregate.

<sup>10</sup> The parameter of 'technology coverage width' is not applicable for the technical categories as it will lead to circular analysis.

Table 5.2.2 Summary indicators of patent activities by 32 Biotechnology sub-categories, 2005-2014<sup>11</sup>

Biotechnology subcategories		Total Patent Families	CAGR (2005-2013)	Avg. Patent Remaining Life	Avg. Time-weighted Patent Citation	Grant Success Rate	Avg. Geo-coverage of Patent Application
Agricultural	Animal Breeding	323	-12%	13.8	1.0	39%	3
	Animal Transgenesis	461	-16%	13.3	1.0	41%	3
	Bioproduct Production	14,171	24%	16.1	0.1	29%	1
	Plant Breeding	2,923	4%	14.8	0.5	40%	2
	Plant Protection	1,267	-6%	14.1	0.7	42%	3
	Plant Transgenesis	3,121	-6%	14.0	0.6	44%	2
Industrial, Food, Environmental	Biobased Fuels & Chemicals	24,284	16%	15.2	0.5	54%	2
	Biopolymers	1,160	6%	14.8	0.6	48%	2
	Biorecycling	1,407	24%	16.0	0.3	49%	1
	Bioremediation	1,240	16%	15.5	0.3	48%	2
	Biosensors	427	16%	15.3	0.4	45%	2
	Enzymes	79,694	5%	14.8	0.6	39%	2
	Fermentation and Bioreactors	905	19%	15.6	0.3	48%	1
	Food Biotechnology	4,534	4%	14.9	0.6	38%	2
	Food Quality and Safety	19,075	16%	15.7	0.3	45%	1
	Food Safety Assessment	9,189	13%	15.1	0.4	47%	2
	Microbial Biodiversity	4,279	14%	15.5	0.5	47%	2
	Microbial Enhanced Production	321	25%	15.1	0.7	49%	2
	Microbial Fermentation Processes	3,351	12%	15.1	0.3	49%	2
Medical, Clinical, Pharmaceutical	Biopharming	277	-8%	13.8	0.6	43%	3
	Cancer	9,121	-13%	13.7	1.3	33%	4
	Cardiovascular Diseases	3,916	-17%	13.6	2.4	36%	4
	Chronic Diseases	14,811	-15%	13.6	1.0	35%	4
	Diabetes	7,537	-18%	13.6	1.3	34%	4
	Gene Therapy	3,639	-14%	13.6	2.1	35%	3
	Growth and Old Age Diseases	3,960	-21%	13.4	1.4	33%	4
	Immunology and Immunotherapy	5,586	-20%	13.3	1.4	35%	4
	Infectious Diseases	24,278	-6%	14.1	0.7	37%	3
	Proteomics and Biomarkers	5,749	-1%	14.4	0.7	40%	3
	Stem Cell Technology	4,538	-1%	14.4	2.1	38%	3
	Stem Cell Therapy	1,041	-16%	13.6	1.8	31%	4
	Toxicology	706	-14%	13.6	2.6	39%	4

<sup>11</sup> The cells are highlighted to reflect the distinction between the listed indicators across 32 subcategories. Overlaps between subcategories are to be expected due to the definition and methodology of categorization mentioned in Section 3.2. Overlaps between subcategories are to be expected due to the definition and methodology of categorization mentioned in Section 3.2. This also holds true for the following Table 5.1.3 and Table 5.1.4.

Figure 5.2.1 shows that two-thirds of Biotechnology patent publications of developing countries were in Industrial, Food, and Environmental Biotechnology, and this sub-field has been keeping growing rapidly in the last ten years. Further investigations need to be conducted before drawing any conclusions on the decline of Medical, Clinical, Pharmaceutical Biotechnology in later years.

It can be seen from Table 5.2.2 that, Enzymes (79,694) records the highest number of patent families among all the 32 sub-categories, and is followed by Biobased Fuels & Chemicals (24,284), and Infectious Diseases (24,278). Among the categories, in the Medical, Clinical, Pharmaceutical biotechnology field, the highest number of patent publications are filed in Infectious Diseases. In Agricultural biotechnology field, Bioproduct Production leads the area by recording the highest number of patent families. In the Industrial, Food, Environmental biotechnology field, the Enzymes is the focus area, which also ranked the highest among all the 32 categories.

CAGR is negative for many of the sub-categories, which means that over the years there has been a decline in the filing trend for these sub-disciplines. However, there are quite a few sub-categories, which have shown a substantially positive growth. The highest CAGR of 25% is shown by Microbial Enhanced Production; Bioproduct Production and Biorecycling with each of them recording a CAGR of 24% follow closely. Fermentation and Bioreactors is next in line in terms of growth, which records a 19% in CAGR.

The average patent remaining life recorded for all the sub-categories falls between 13 to 15 years, which means that most of the filings in the technology are 5 to 7 year old. Bioproduct Production, Microbial Biodiversity, Food Quality and Safety, Fermentation and Bioreactors, and Biorecycling all score 16 on average patent life index.

The average time weighted patent citation, which is the measurement for the impact a particular invention or patent family, is having within its technical field is recorded highest for Toxicology. Toxicology is closely followed by Cardiovascular Diseases with 2.40, Gene Therapy with 2.15 and Stem Cell Therapy with 2.14 as the average time weighted patent citation. Most of the other sub-categories fall under less than 1 in the citation impact parameter index.

The average grant success rate is highest for Biobased Fuels & Chemicals, which is more than 50%. Biobased Fuels & Chemicals is closely followed by Biorecycling, Bioremediation, Fermentation and Bioreactors, Food Safety Assessment, Food Quality and Safety, which all record a more than 45% grant success rate. The lowest rate recorded is 29% for Bioproduct Production.

The average geographical coverage of patent applications is highest for Toxicology, Stem Cell Therapy, Immunology and Immunotherapy, Growth & Old Age Diseases, Diabetes, Chronic Disease, Cardiovascular Disease and Cancer. Evidently, geographical coverage is highest for disease related sub-categories. The lowest average geographical coverage of one is recorded for Bioproduct Production, Food Quality and Safety, Fermentation and Bioreactors, and Biorecycling.

Table 5.2.3 Total patent families by 32 Biotechnology sub-categories for each region, 2005-2014<sup>12</sup>

Biotechnology subcategories		Latin America and the Caribbean Region	East, South-East Asia and the Pacific Region	Arab Region	Central and South Asia Region	Sub-Saharan Africa Region	S&T Lagging Region
Agricultural	Animal Breeding	99	241		23	26	3
	Animal Transgenesis	120	356	1	78	31	2
	Bioproduct Production	296	13,989	10	130	23	45
	Plant Breeding	1,081	2,417	18	302	150	67
	Plant Protection	723	889	13	188	107	34
	Plant Transgenesis	1,290	2,313	14	467	204	45
Industrial/Food/Environmental	Biobased Fuels & Chemicals	3,333	21,438	65	1,443	345	290
	Biopolymers	243	1,035	2	108	29	13
	Biorecycling	75	1,365	3	32	7	14
	Bioremediation	145	1,126	2	70	24	17
	Biosensors	27	396	1	29	4	
	Enzymes	16,687	68,992	402	7,452	2,433	1,724
	Fermentation and Bioreactors	33	876		24	3	1
	Food Biotechnology	1,201	3,783	16	413	195	45
	Food Quality and Safety	1,653	17,824	22	667	198	104
	Food Safety Assessment	742	8,395	8	559	96	29
	Microbial Biodiversity	590	3,883	13	270	97	56
Medical/Clinical/Pharmaceutical	Microbial Enhanced Production	101	276	1	17	7	4
	Microbial Fermentation Processes	288	3,181	4	134	35	25
	Biopharming	135	221	3	44	20	17
	Cancer	3,900	7,372	155	1,410	565	532
	Cardiovascular Diseases	1,688	3,099	52	664	240	211
	Chronic Diseases	6,882	11,613	255	2,442	1,029	916
	Diabetes	3,586	5,735	119	1,363	520	489
	Gene Therapy	1,169	2,804	32	627	197	145
	Growth and Old Age Diseases	1,949	3,060	67	701	264	303
	Immunology and Immunotherapy	2,831	4,148	85	1,061	413	356
	Infectious Diseases	7,278	19,359	204	3,492	1,315	772
	Proteomics and Biomarkers	1,067	4,981	22	654	159	74
	Stem Cell Technology	824	3,857	15	556	151	55
	Stem Cell Therapy	239	815	2	187	40	11
	Toxicology	452	587	6	115	38	83

<sup>12</sup> The cells are highlighted to reflect the relative research focus in each region.



Table 5.2.4 Total patent families by 32 Biotechnology sub-categories for top 10 countries/territories, 2005-2014<sup>13</sup>

Biotechnology subcategories		Argentina	Brazil	China	Colombia	India	Mexico	Philippine s	South Africa	Taiwan, China	Vietnam
Agricultural	Animal Breeding	5	34	230	1	23	54		26	11	
	Animal Transgenesis	15	28	322	4	78	67	4	31	22	6
	Bioproduct Production	36	96	13,874	14	130	98	5	23	73	8
	Plant Breeding	273	148	2,185	68	302	398	70	150	80	75
	Plant Protection	186	127	761	40	186	253	40	107	29	49
	Plant Transgenesis	348	239	2,101	49	465	483	64	204	63	74
Industrial/Food/Environmental	Biobased Fuels & Chemicals	403	1,403	20,337	139	1,431	1,070	139	345	496	200
	Biopolymers	31	76	960	13	108	98	9	29	40	18
	Biorecycling	8	20	1,321	4	30	32	8	7	17	8
	Bioremediation	17	34	1,087	6	70	60	13	24	21	3
	Biosensors	1	13	365	1	29	11		4	30	1
	Enzymes	2,399	3,053	63,238	798	7,434	7,318	1,048	2,433	3,233	918
	Fermentation and Bioreactors	3	17	859	1	24	11	1	3	12	4
	Food Biotechnology	291	255	3,547	27	412	495	77	195	67	74
	Food Quality and Safety	208	512	17,139	43	660	744	111	198	418	84
	Food Safety Assessment	74	247	7,994	22	557	342	36	96	313	28
	Microbial Biodiversity	71	144	3,677	28	269	271	33	97	101	39
	Microbial Enhanced Production	11	17	264	15	16	50	1	7	4	2
	Microbial Fermentation Processes	31	91	3,041	6	134	125	18	35	77	29
Medical/Clinical/Pharmaceutical	Biopharming	24	21	185	7	44	48	4	20	21	9
	Cancer	466	467	5,905	236	1,407	1,751	279	565	866	244
	Cardiovascular Diseases	218	191	2,509	105	660	749	118	240	347	96
	Chronic Diseases	826	795	9,125	421	2,439	3,102	530	1,029	1,392	433
	Diabetes	452	359	4,457	221	1,359	1,644	275	520	708	227
	Gene Therapy	120	158	2,442	60	625	604	78	197	207	57
	Growth and Old Age Diseases	225	194	2,388	124	700	863	136	264	375	129
	Immunology and Immunotherapy	289	387	3,229	166	1,057	1,326	177	413	526	169
	Infectious Diseases	820	1,110	16,711	419	3,486	3,465	562	1,315	1,421	497
	Proteomics and Biomarkers	100	191	4,474	40	654	588	48	159	419	26
	Stem Cell Technology	70	162	3,452	35	554	445	51	151	305	39
	Stem Cell Therapy	25	39	722	8	186	141	19	40	61	8
	Toxicology	49	51	444	31	115	174	19	38	81	39

<sup>13</sup> The cells are highlighted to reflect the relative patent focus in each country/territory.

### 5.3 Research Publications v.s. Patent Activities by Sub-Categories

---

By comparing the major findings from the sub-category analyses on research publications and patent activities, it can be seen that:

- The academic research has been focusing more on Medical, Clinical, Pharmaceutical Biotechnology, while technologies were most active in Industrial, Food, and Environmental Biotechnology.
- Infectious Diseases, Proteomics and Biomarkers, and Chronic Diseases are the top 3 most productive sub-categories in scientific publishing. And Infectious Diseases accounted for more than one fourth of the total publication in overall biotechnology field.
- Enzymes, Biobased Fuels & Chemicals and Infectious Diseases are the top 3 sub-categories in which patent filings were most active. Furthermore, over half of the patents in Biotechnology field are categorized under Enzyme, which indicates the strong technology focus in this area.
- The Compound Annual Growth Rate of patent filings shows that Microbial Fermentation Processes, Biorecycling and Bioproduct Production are the top areas with highest increase, indicating huge technology/business opportunities underlying these sub-categories.

## 6 Overview of Biotechnology by Key Metrics

### 6.1 Research Publication Overview

---

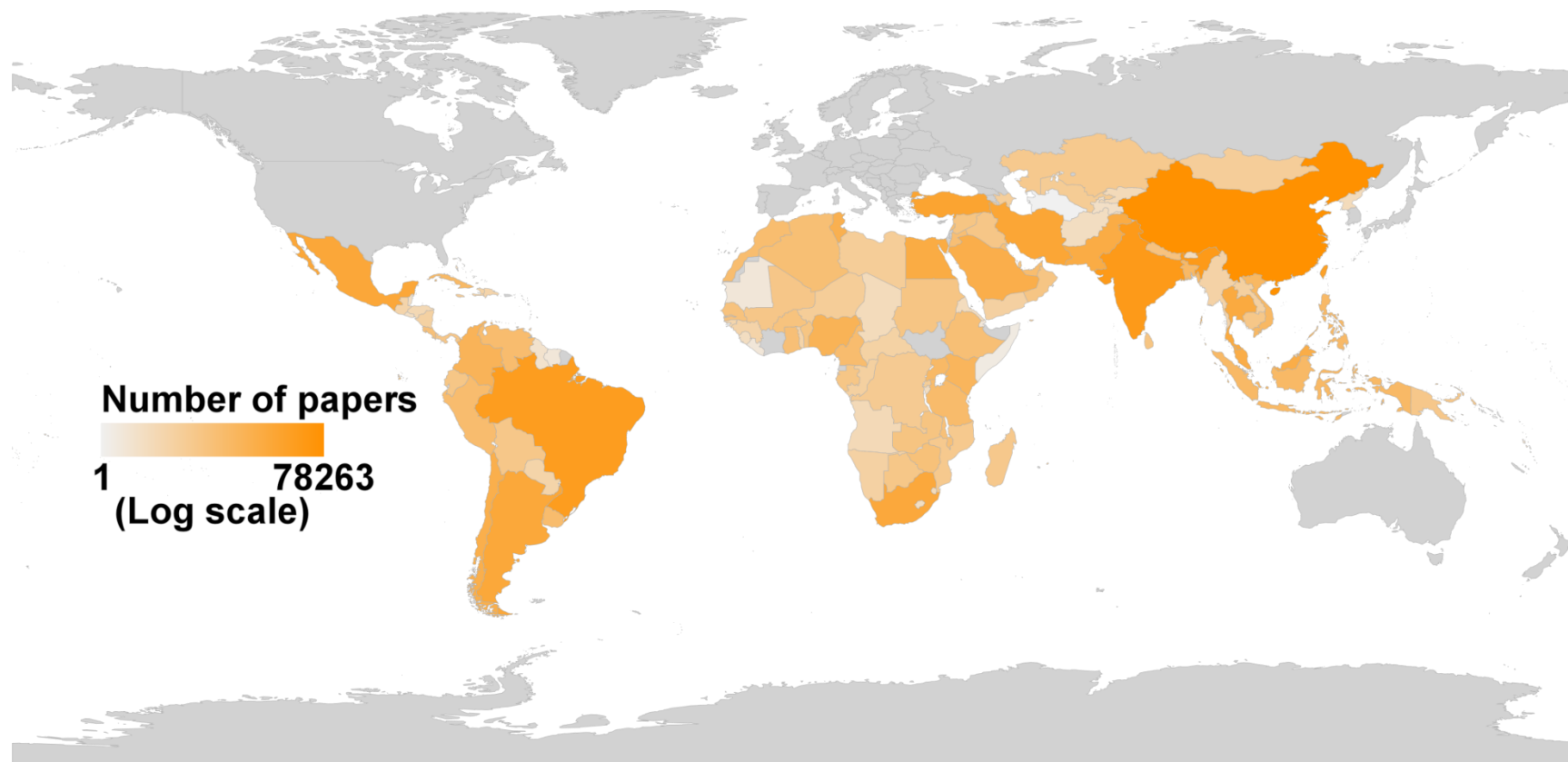


Figure 6.1.1 Overview by number of papers, 2005-2014

Figure 6.1.1 shows research papers in Biotechnology by country between 2005 and 2014. The number of publications from shown countries is indicated by intensity of orange colour; those that appear darker have published in this area more frequently than those in lighter orange. China had the highest number of papers (78,263), followed by India (24,081) and Brazil (17,769).

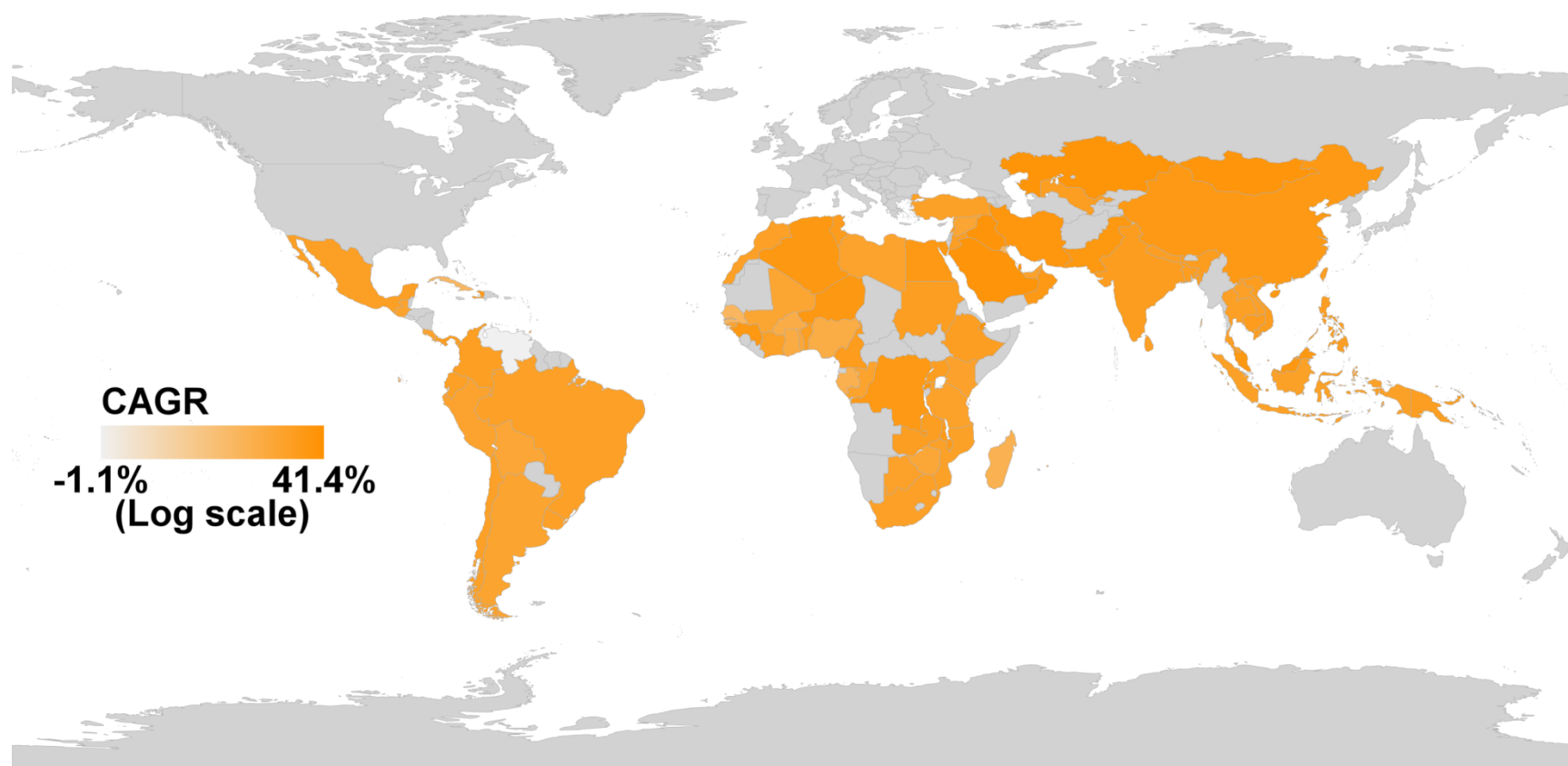


Figure 6.1.2 Overview by CAGR of publication output, 2005-2014

Figure 6.1.2 shows the Compound Annual Growth Rate (CAGR) of publications for developing countries in Biotechnology. The CAGR was calculated only for countries which had publications for a continuous set of years. For these countries the start year was the earliest year of publication, and the end year was the most recent publication year up to 2014. The CAGR was not calculated for those countries which had sporadic (dis-continuous) publications. Djibouti had the highest CAGR (41.4%) and Venezuela had the lowest (-1.1%), which is also the only one negative growth rate.

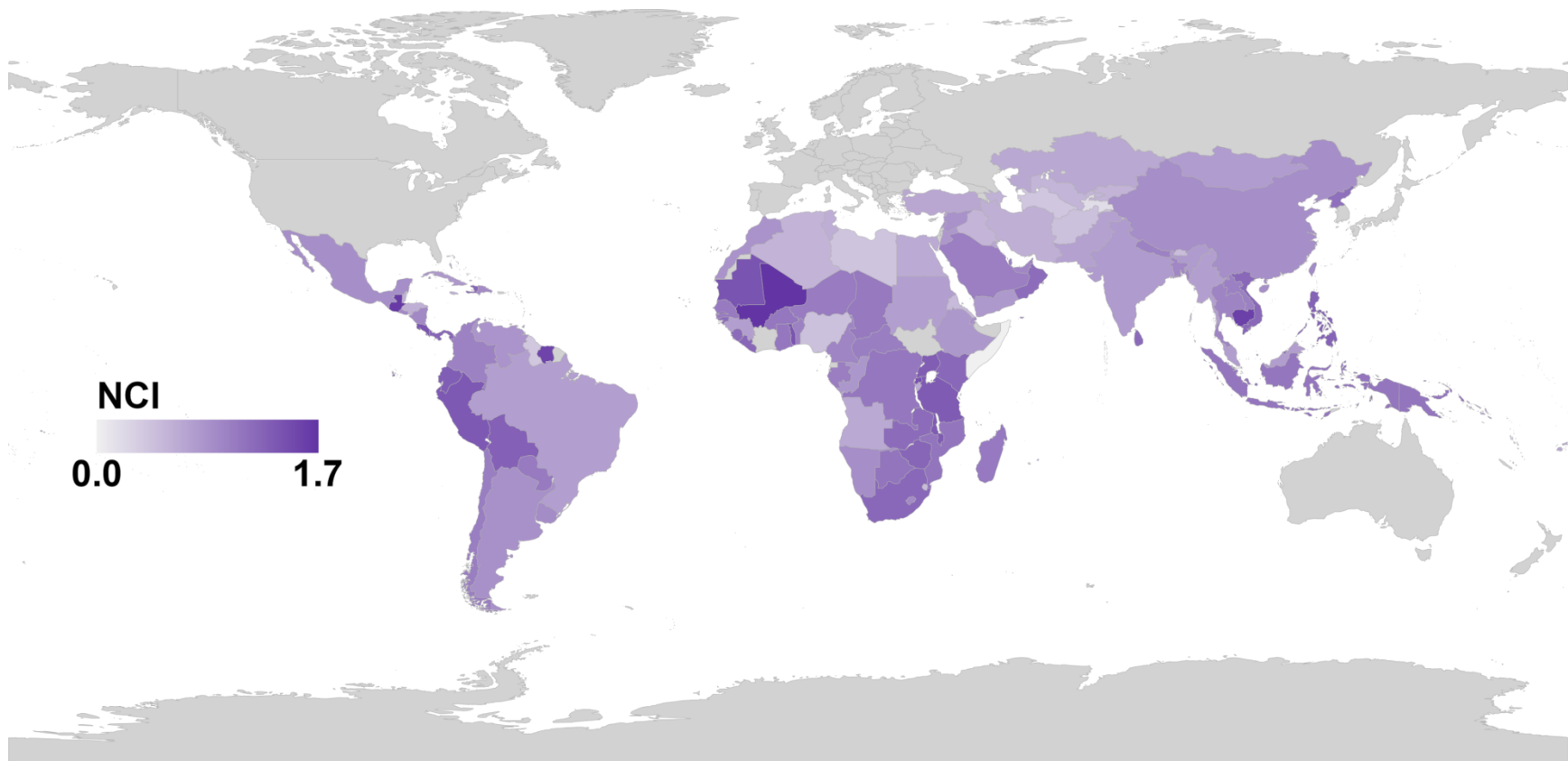


Figure 6.1.3 Overview by normalized citation impact, 2005-2014

This figure shows the normalized citation impact of the papers published by each country. It should be noted that this indicator can be skewed by small paper counts. For example Suriname had the third highest citation impact (1.56) of all the countries, but published three papers between 2005 and 2014.

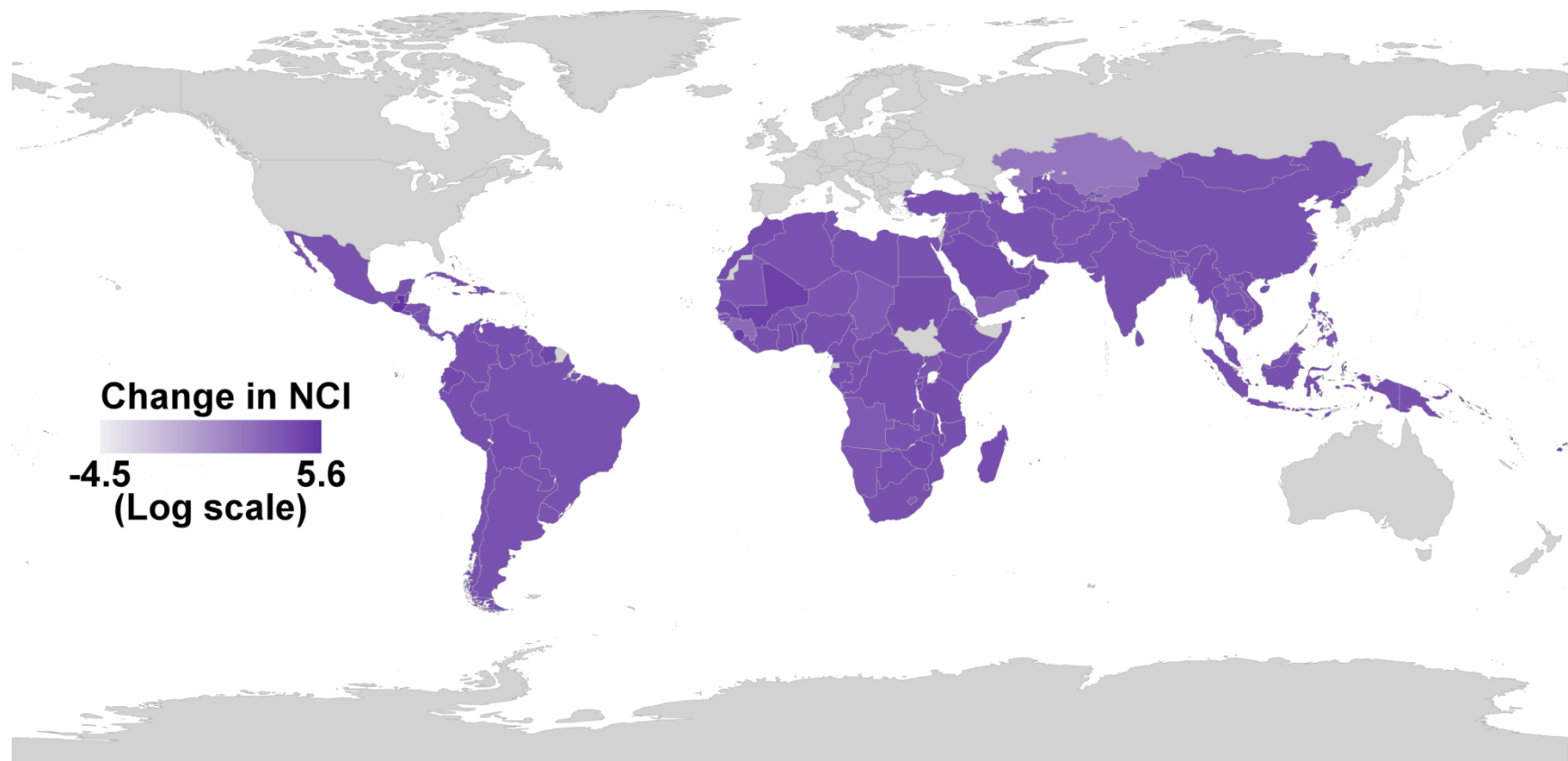


Figure 6.1.4 Overview by changes in normalized citation impact, 2005-2014

Figure 6.1.4 shows the change in the normalized citation impact (NCI) between 2005 and 2014. This change is defined as the difference in the average citation impact of all papers from that country between 2005 and 2014. The majority of the countries had positive changes in their NCI. This is most likely indicating that the research across these countries is growing in collaboration which is leading to more citations. The country with the greatest positive change in normalized citation impact was Guatemala.

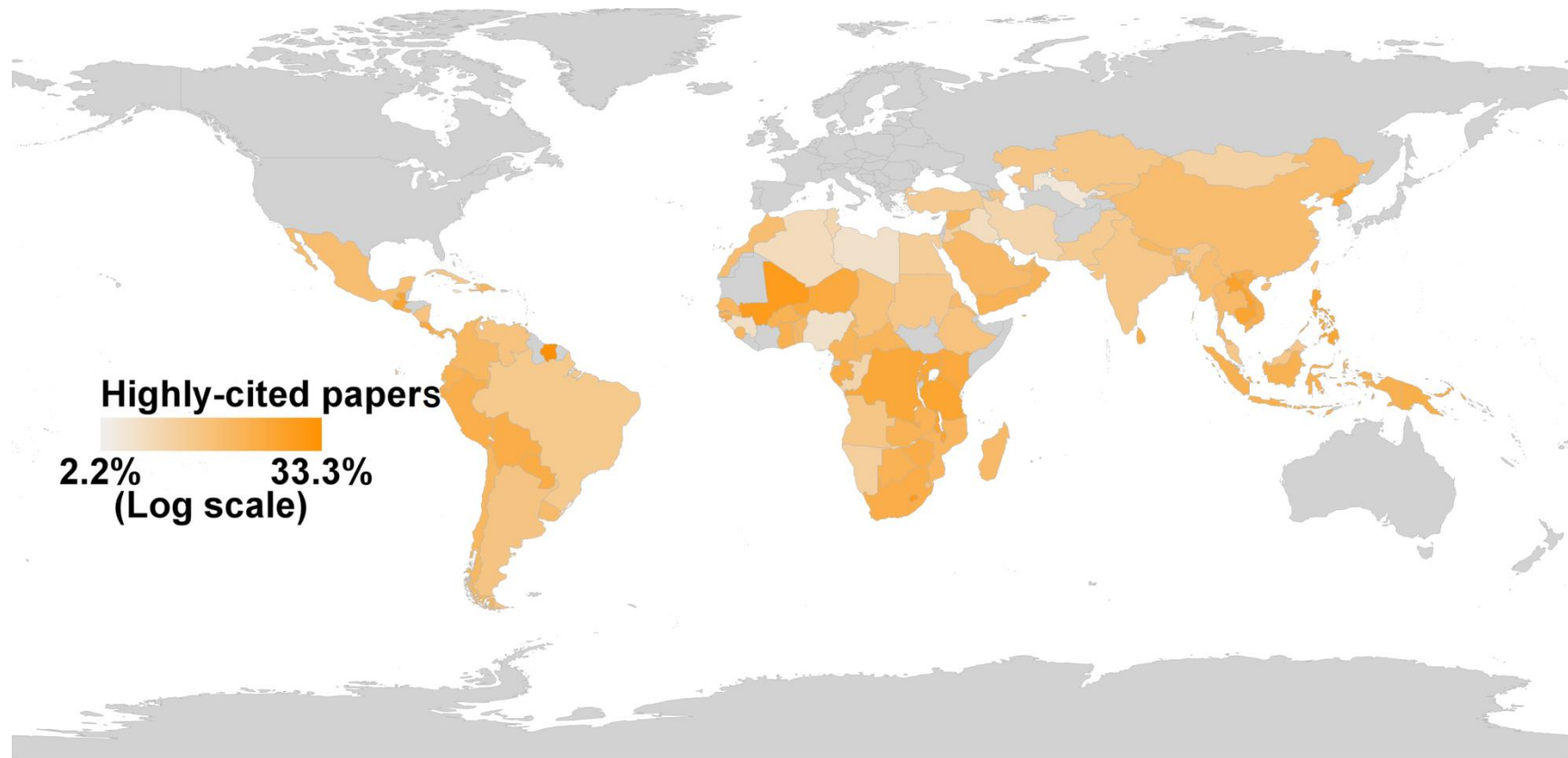


Figure 6.1.5 Overview by highly-cited papers, 2005-2014

Figure 6.1.5 shows which percentage of a country's papers are highly-cited. Highly-cited papers are defined as those that rank in the top 10% of world output, by citation count. Similar to the normalized citation impact this indicator can be skewed by low paper counts. For example the country with the highest percentage of highly-cited papers is Suriname which published three papers between 2005 and 2014. Countries with high percentage of highly-cited papers and large number of papers are: China, Mexico, and South Africa.

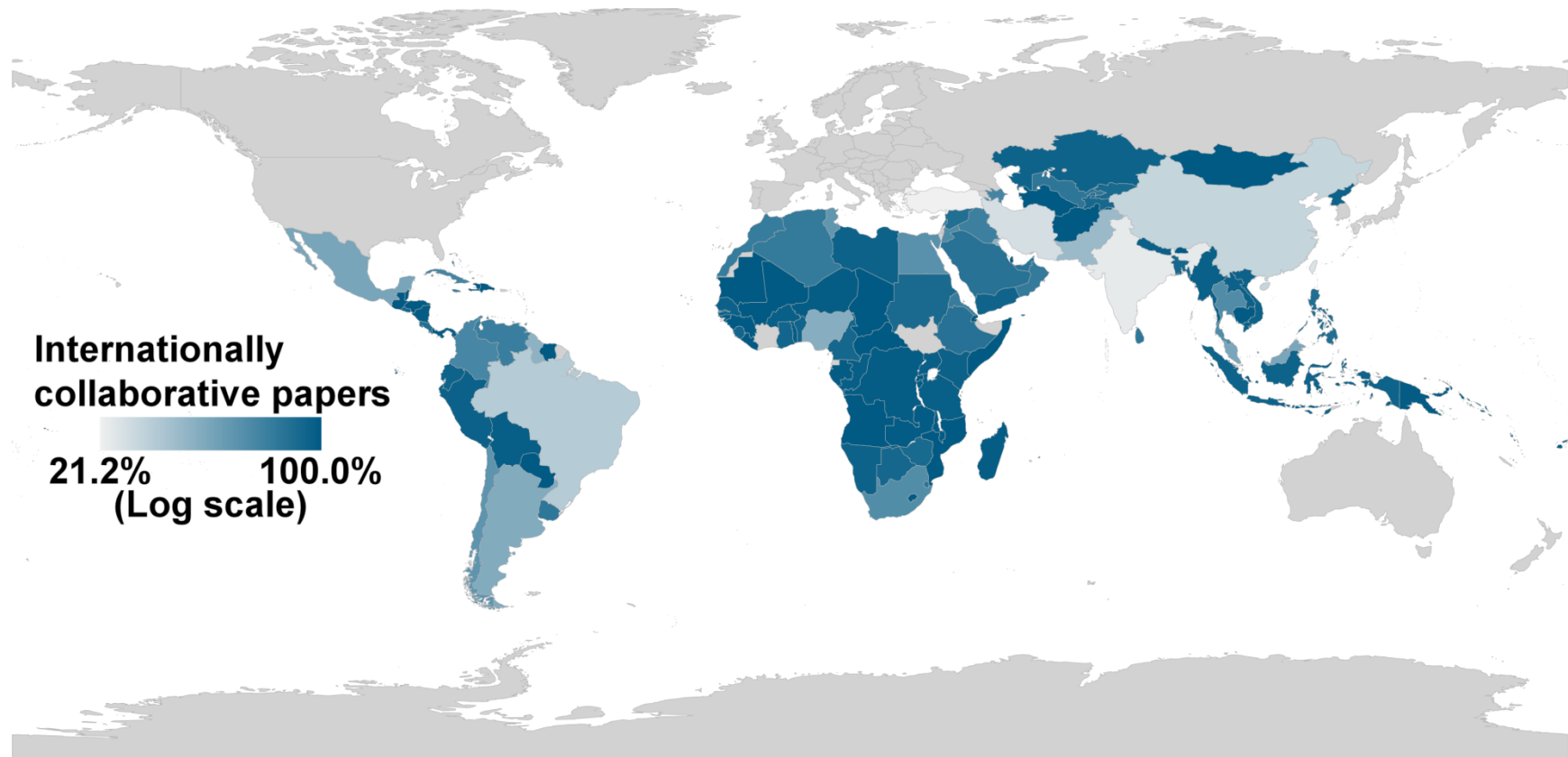


Figure 6.1.6 Overview by internationally collaborative papers, 2005-2014

Figure 6.1.6 shows the percentage of papers which were intentionally collaborative. A paper is internationally collaborative if the author affiliations are from more multiple countries. This figure shows that the geographical region of highest international collaboration is Africa. This is also consistent with the findings in Figure 4.1.7.



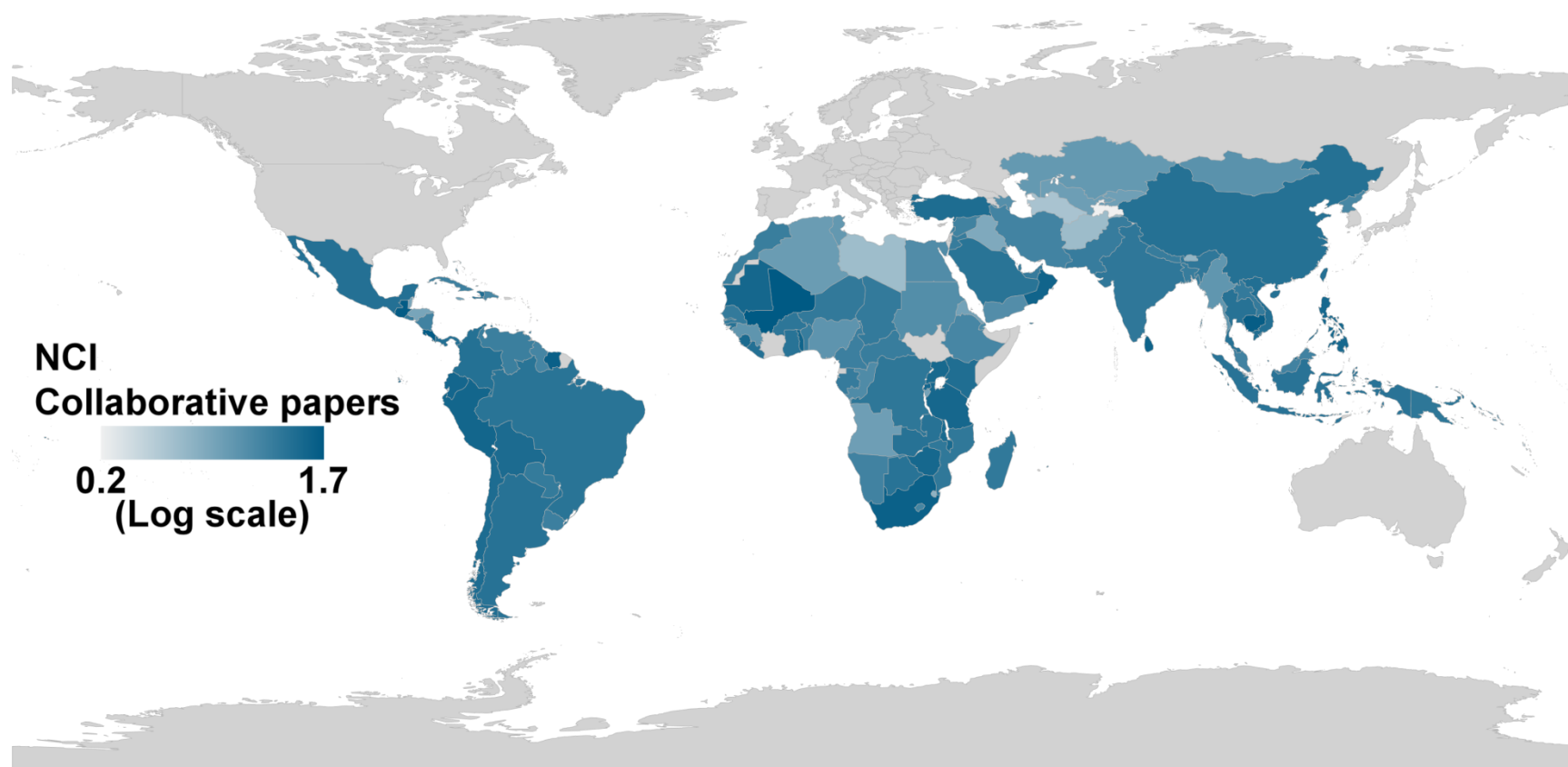


Figure 6.1.7 Overview by normalized citation impact (NCI) of internationally collaborative papers, 2005-2014

Figure 6.1.7 shows the normalized citation impact of the internationally collaborative papers. The majority of the countries have a normalized citation impact greater than one for their internationally collaborative papers. It should be noted that countries like China and India which have low percentage of international collaborative papers (see Figure 6.1.6), still generate papers with high citation impact.

## 6.2 Patent Activities Overview

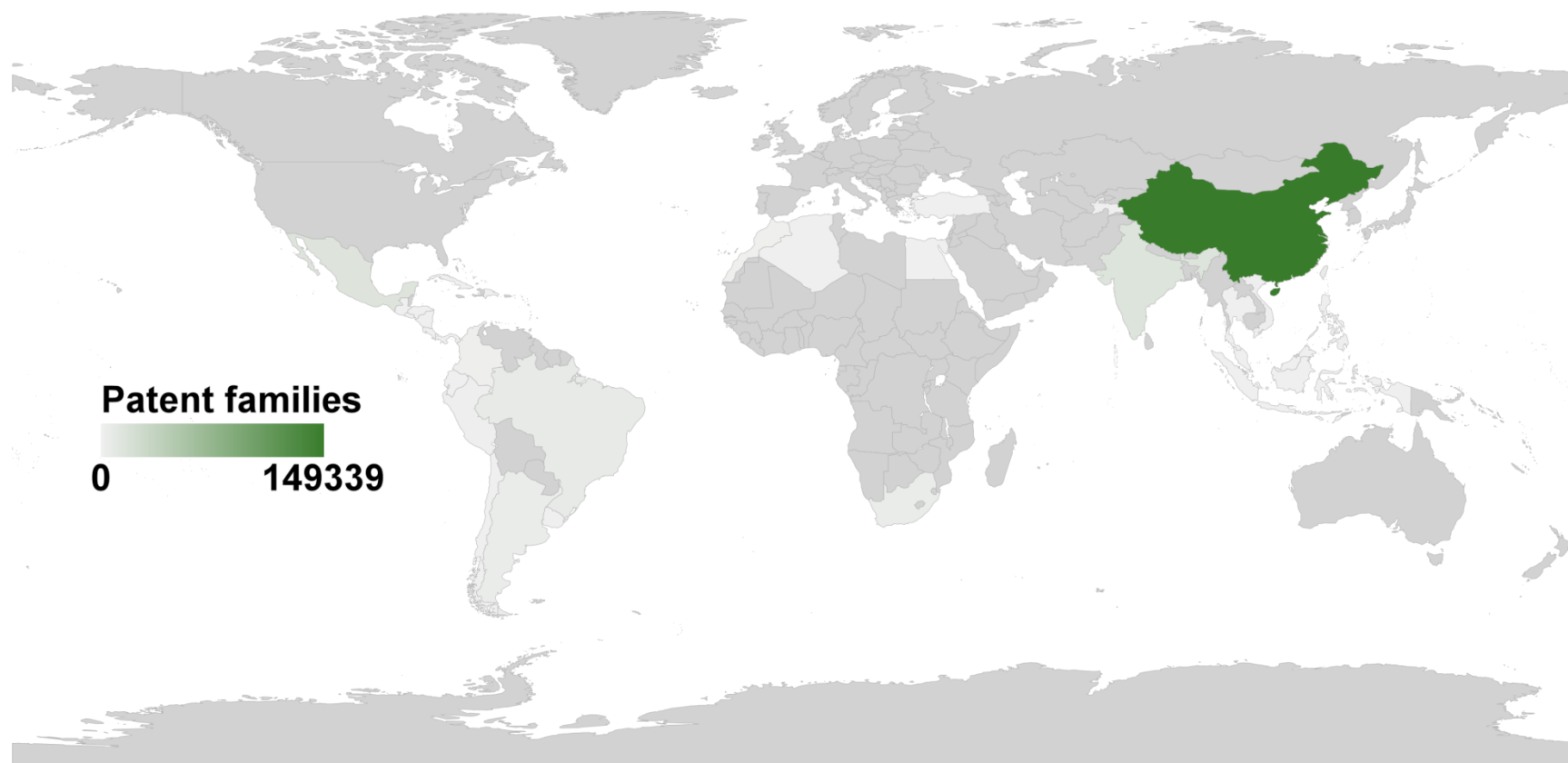


Figure 6.2.1 Overview by total patent families, 2005-2014

As can be seen in Figure 6.2.1, China leads in all countries with 149,339 patent families, followed by India (15,420) and Mexico (14,574). Algeria and Tajikistan ranked the bottom with each recording just a single filing over the period of 10 years.

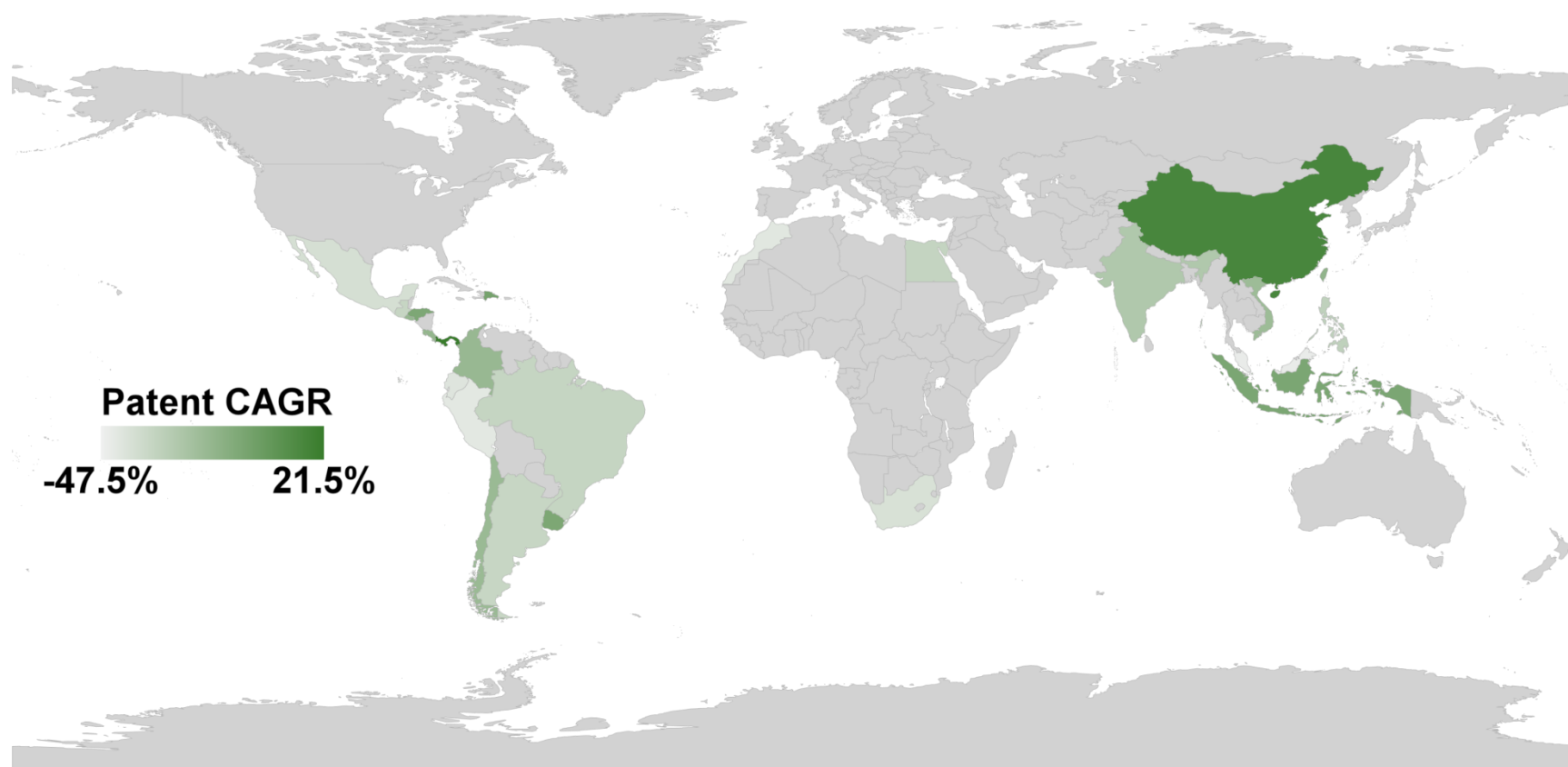


Figure 6.2.2 Overview by CAGR of patent filings, 2005-2013

The CAGR of patent filings in the developing countries between the year 2005 and 2013 shows that China and Panama are the only countries with a positive growth rate among the countries that are a part of the present study. The recorded annual growth is 15.0% for China, and 21.5% for Panama. Since Algeria and Tajikistan have recorded only one filing in the period of this study, CAGR is not applicable for them. As with the publications, the CAGR was calculated only for countries which had patent filings for a continuous set of years. For these countries the start year was the earliest year of patent filing, and the end date was the most recent patent filing year up to 2013. The CAGR was not calculated for those countries which had sporadic (dis-continuous) filing.

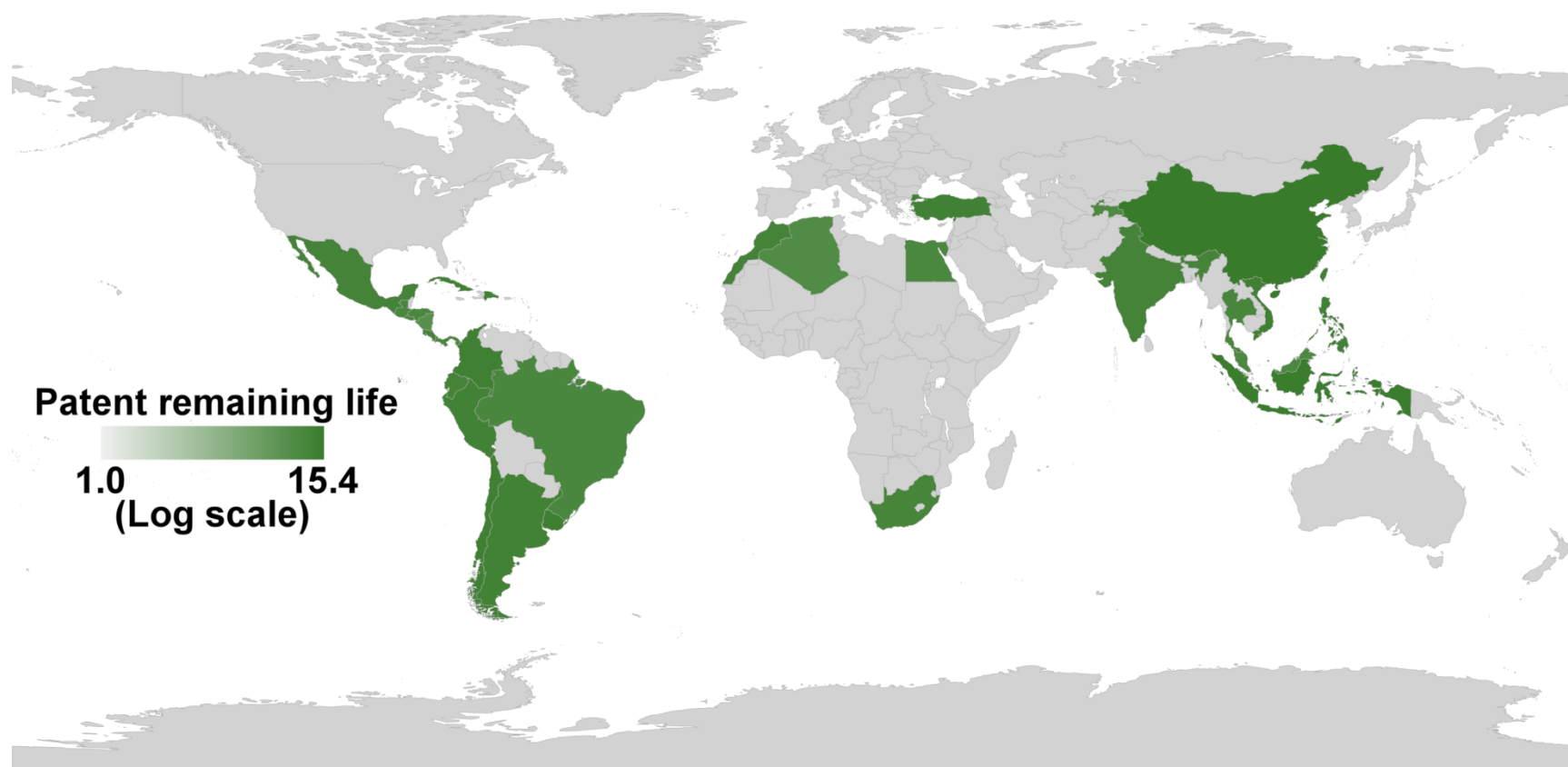


Figure 6.2.3 Overview by average patent remaining life, 2005-2014

The average patent remaining life recorded for all countries falls between 9 to 16 years, which means that most of the filings in the technology are 4 to 11 year old. China records the highest average patent remaining life of 15.41 followed by Indonesia with 14.93.

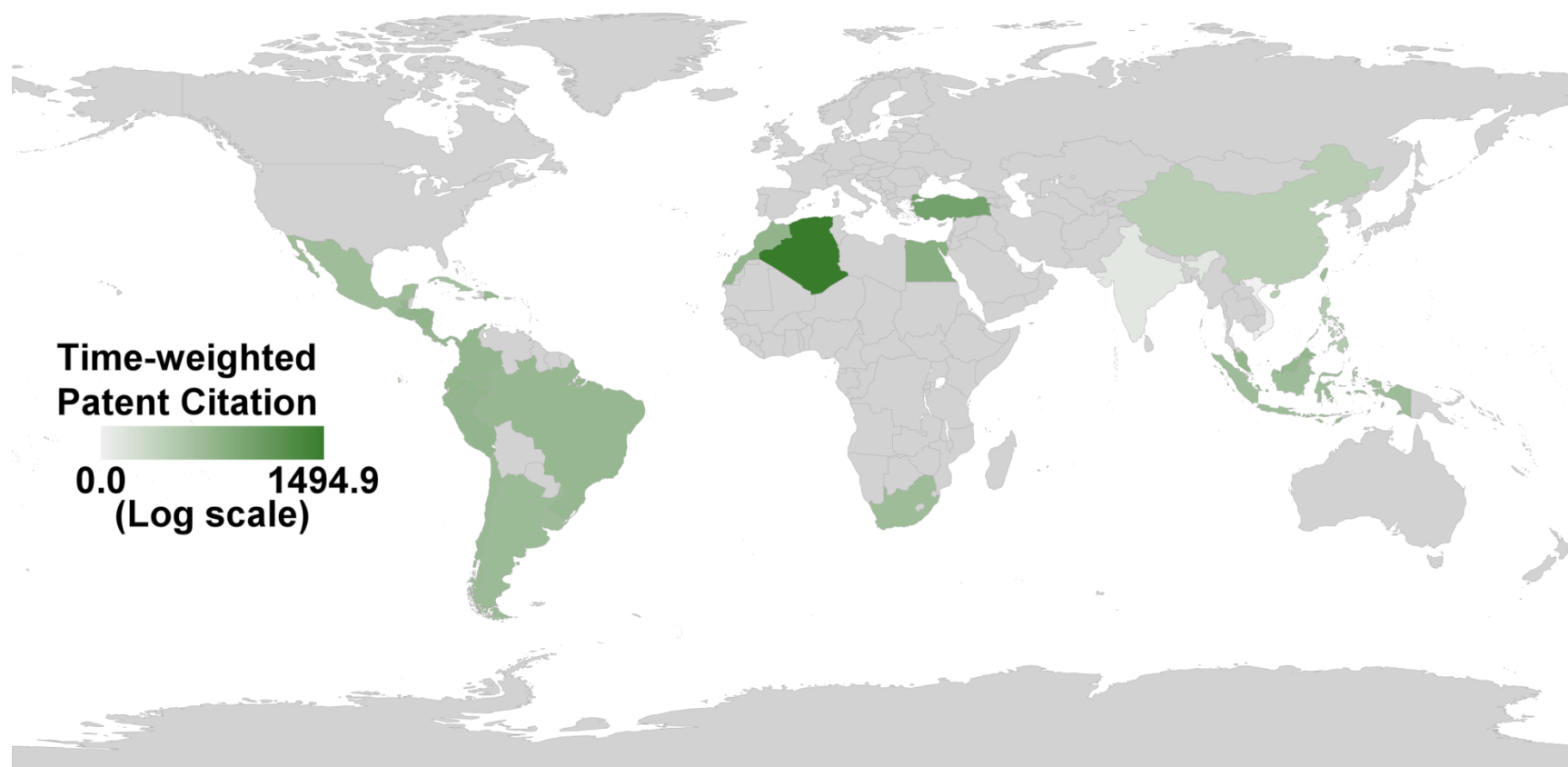


Figure 6.2.4 Overview by average time weighted patent citation, 2005-2014

The average time weighted patent citation, which is the measurement for the impact a particular invention or patent family, is highest for Algeria which 1494.9, however, Algeria has only one filing during the period of this study, thus this does reflect the weightage scenario for this parameter. Algeria is followed by Turkey, which records an average of 38.03. Turkey is followed by Egypt with around 11 and Honduras with 6.46 as its average time weighted patent citation.

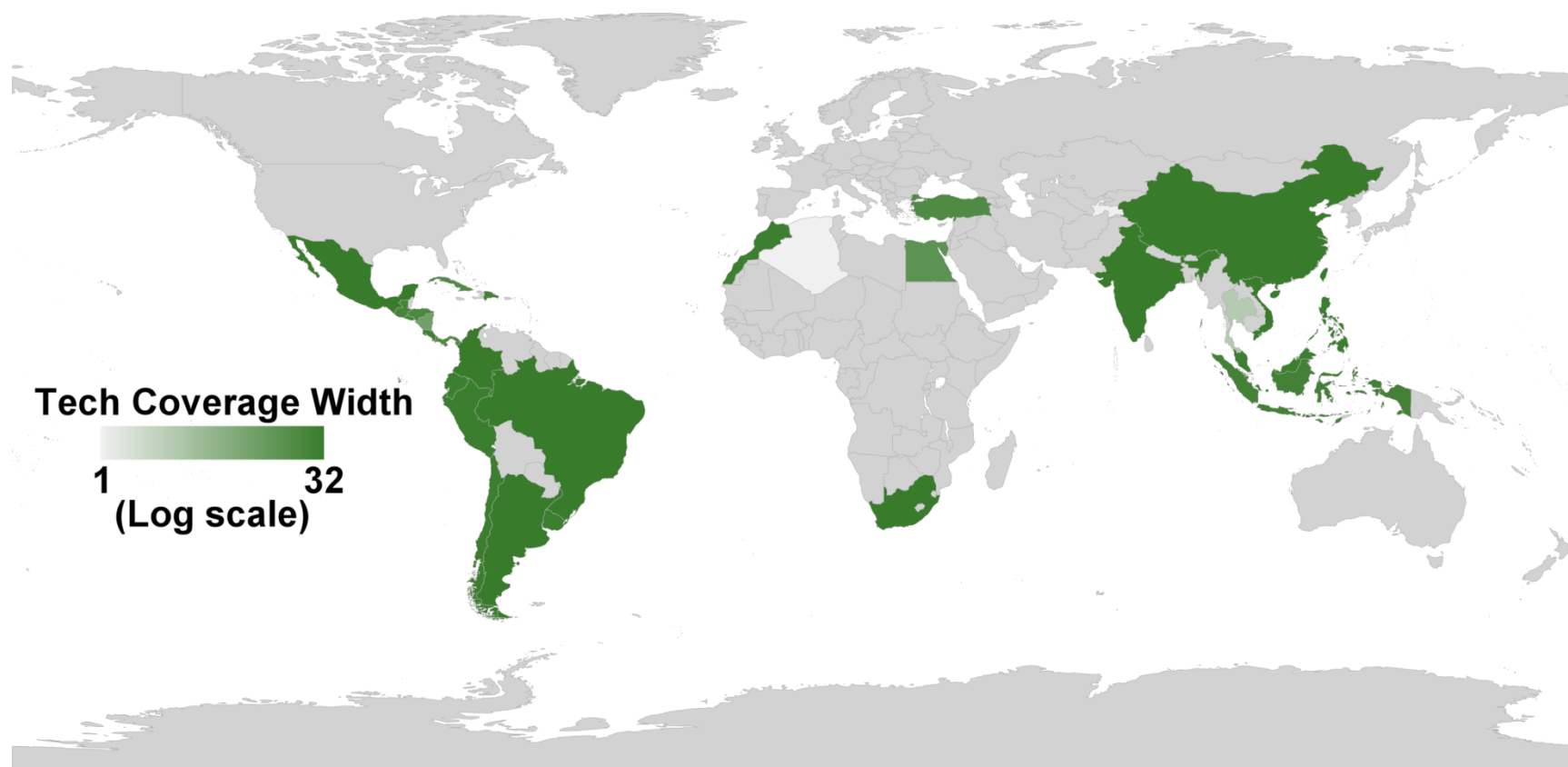


Figure 6.2.5 Overview by average technological coverage width, 2005-2014

The technical coverage width is 32 for China, India, Mexico, Taiwan(China), Brazil, South Africa, Argentina, and Columbia signalling the filing of patents among all the sub-categories of Biotechnology in these countries.

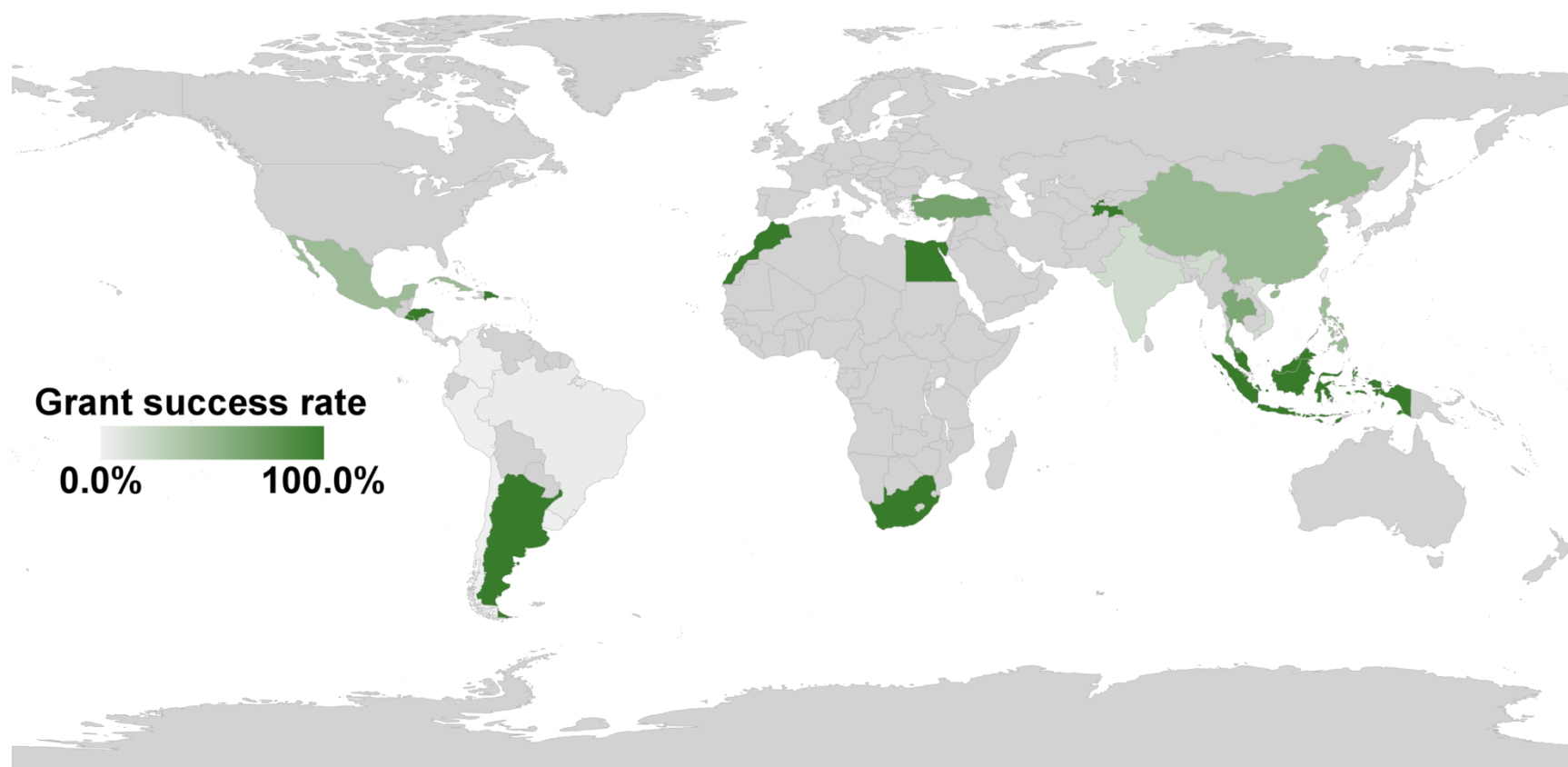


Figure 6.2.6 Overview by average grant success rate, 2005-2014

The average grant success rate is 100% for South Africa along with Argentina, Morocco, Malaysia, Indonesia, and Dominican Republic, and Honduras. However, except South Africa, for all these countries the number of filings is below 100%. China, which has the highest number of filings, has a grant rate of 48%. India, although is second to China in the number of fillings, has only 17% grant rate. For Ecuador, Guatemala, Panama, Nicaragua and Algeria, this parameter could not be ascertained due to the non-availability of grant data which is due to the lack of distinction of published or filed applications from the ones which are granted in these countries.

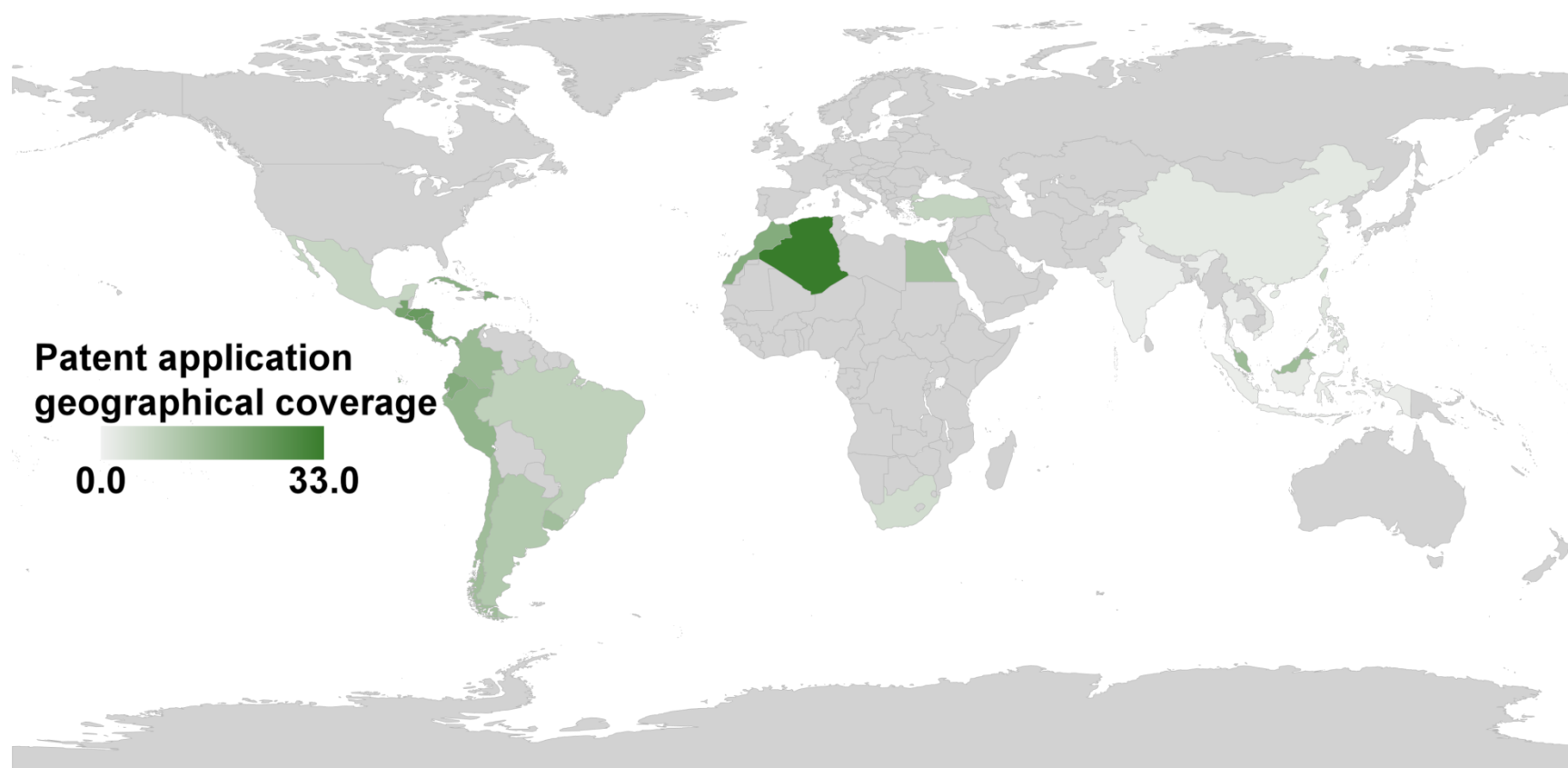


Figure 6.2.7 Overview by average geographical coverage of patent application, 2005-2014

The average geographical coverage of patent applications is highest for Algeria which is 33 and followed by El Salvador with 24 and Nicaragua with 22. However, these countries have a small number of filings. Interestingly, average geographical coverage for China is only two which means that most of the filings are locally filed in China and in one more jurisdiction only.

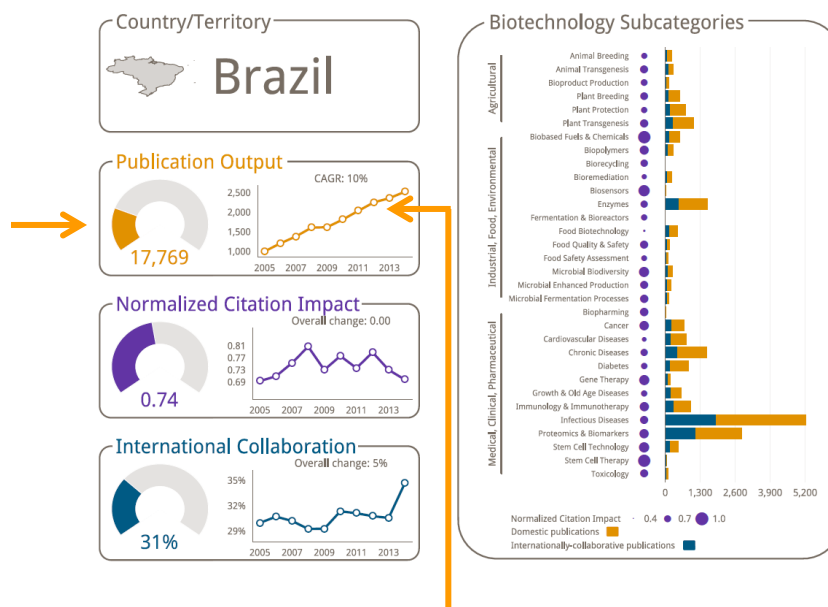


## 7 Profile of Biotechnology by Developing Countries

### 7.1 Publication Profiles

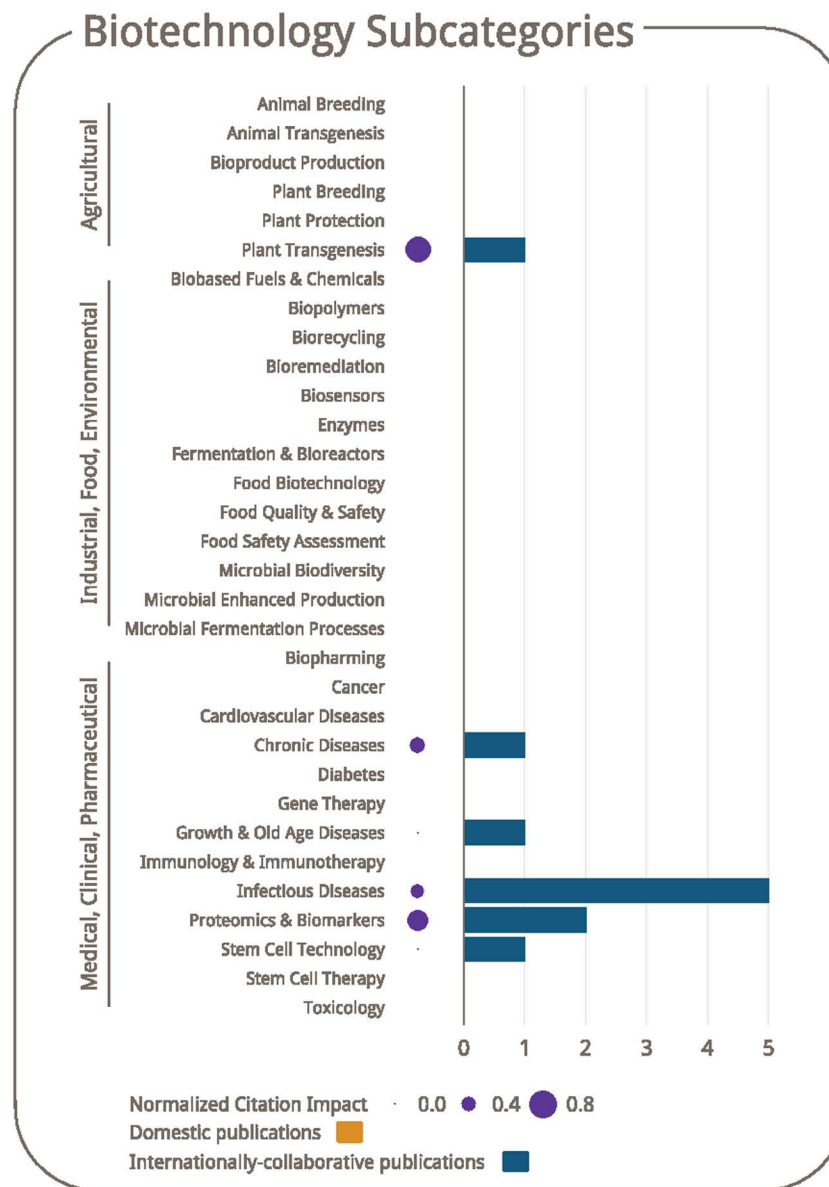
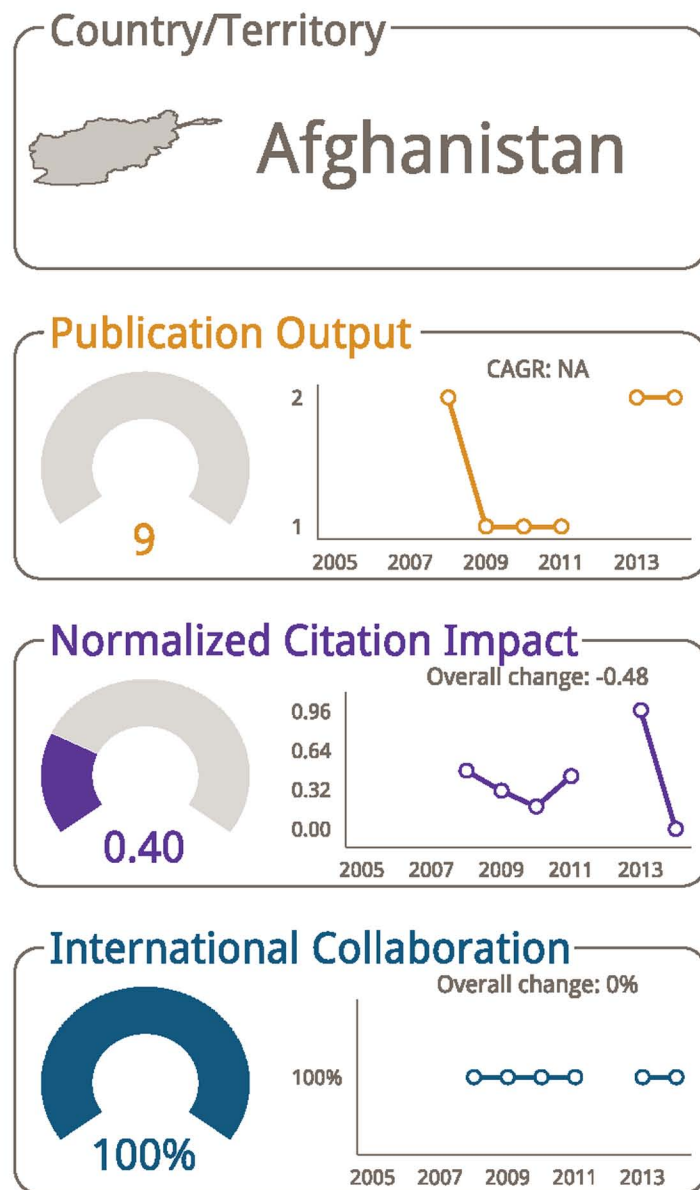
The publications profiles are presented for each of the 128 countries in the following layout.

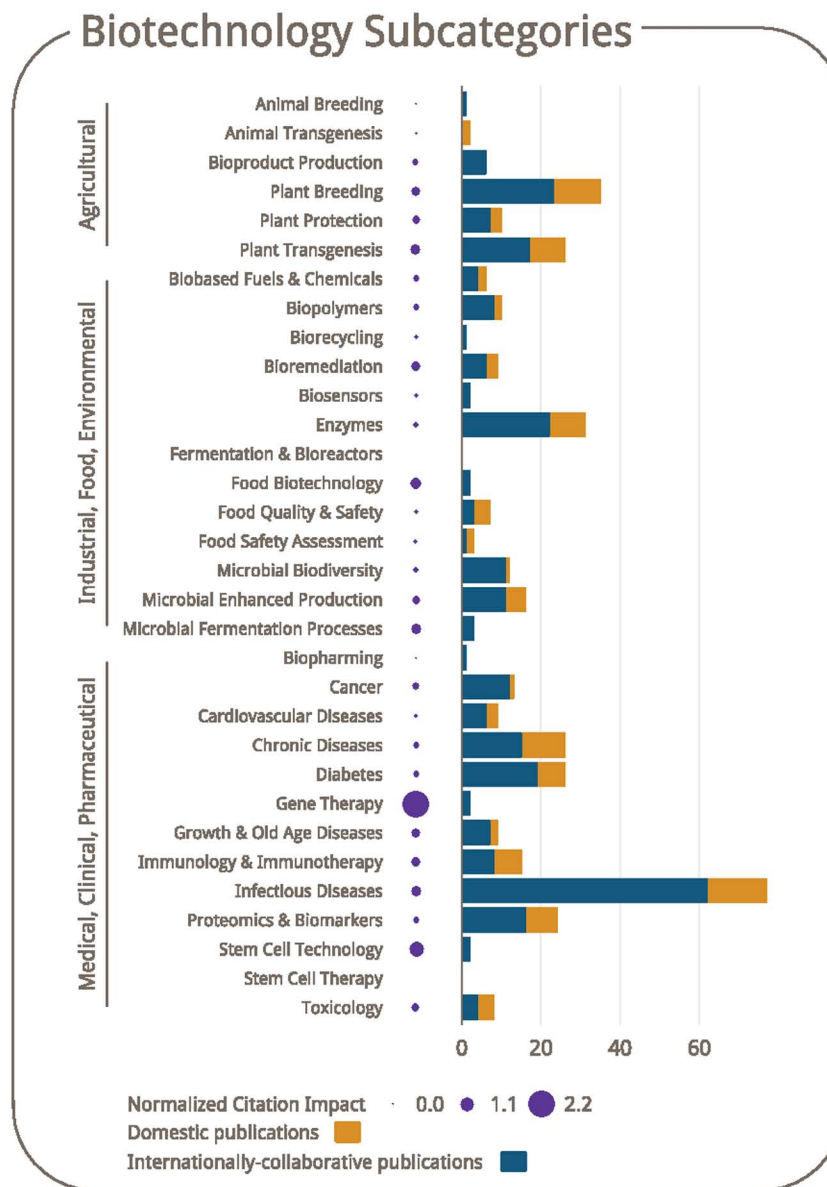
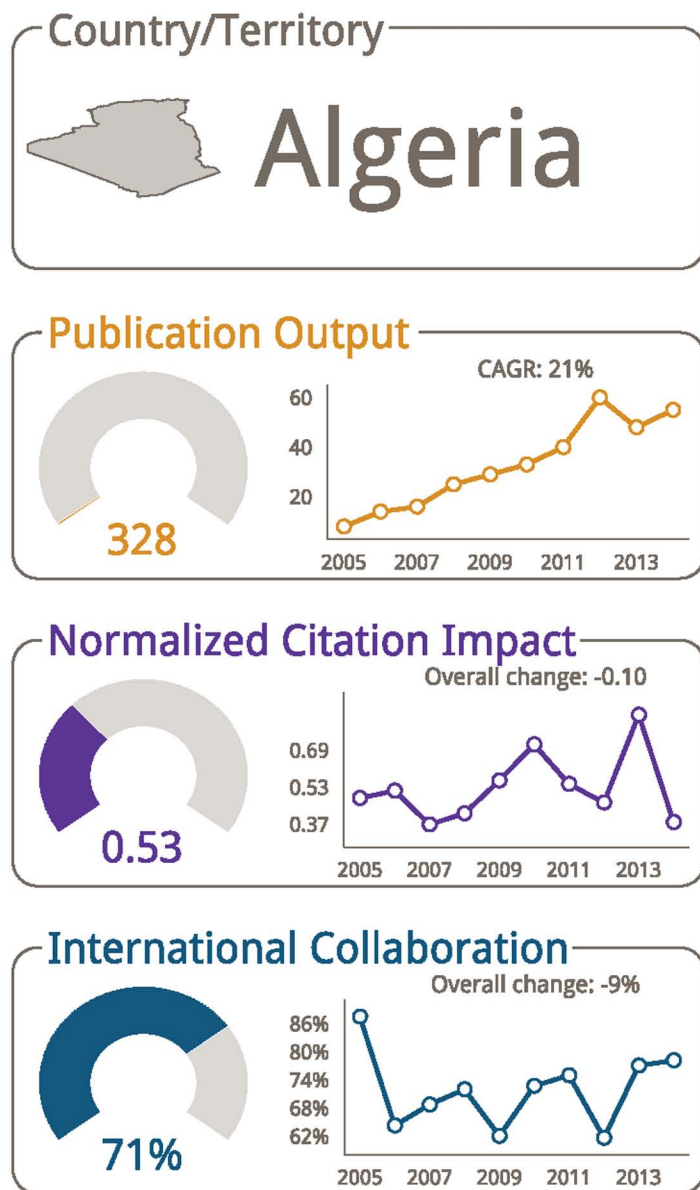
**Dials** show a country's overall level of publication output, Normalized Citation Impact, and international collaboration relative to the minimum and maximum values across all 128 countries. For example the minimum publication output for all the countries is 1 while the maximum is 78,263 (China). Brazil's publication output of 17,769 is about one quarter of the way from the minimum value to the maximum value, and hence its dial is about a quarter of the way over.

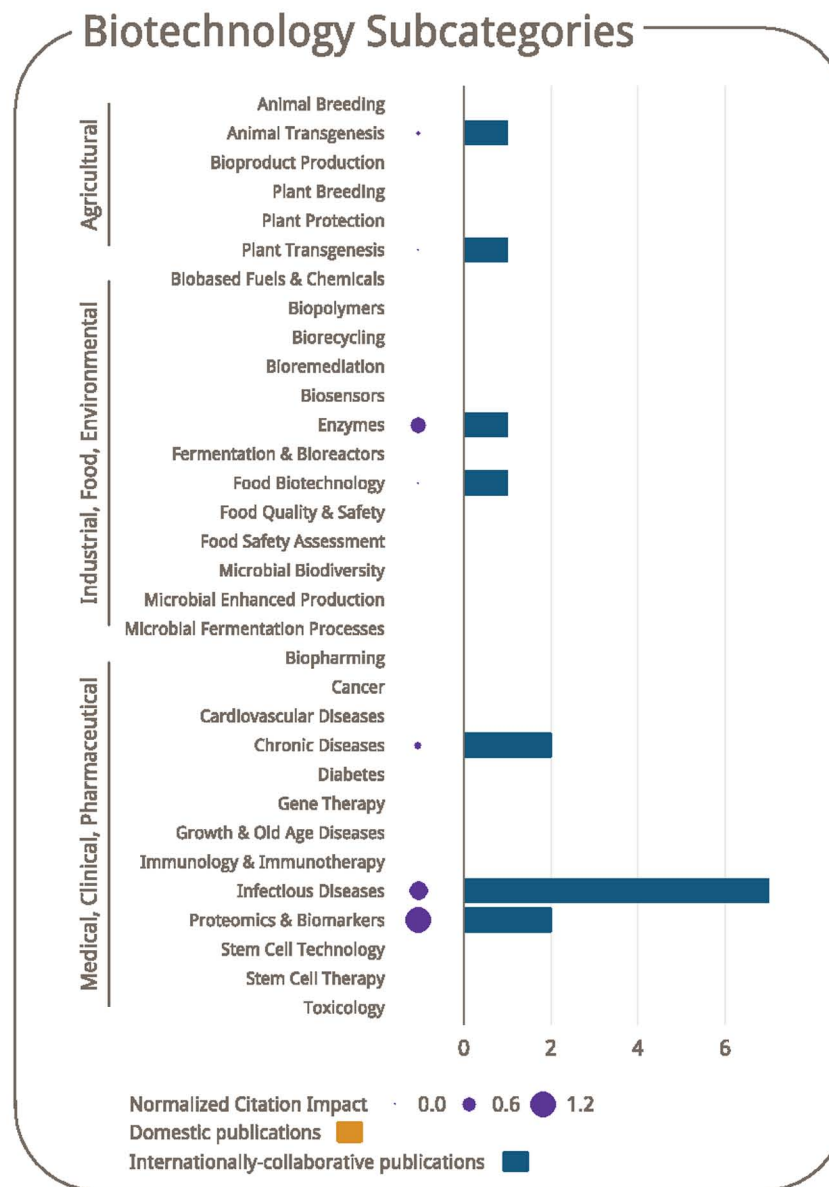
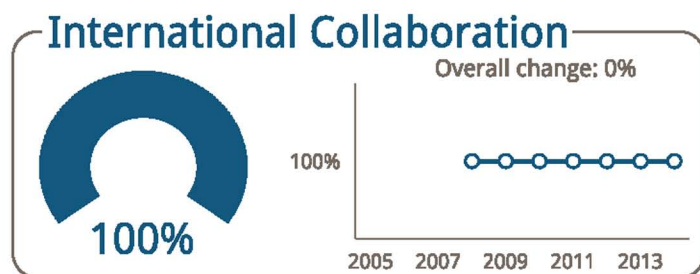
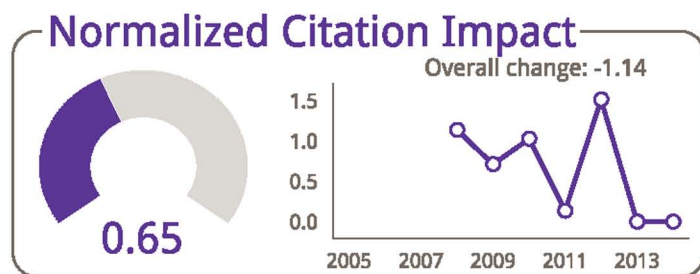
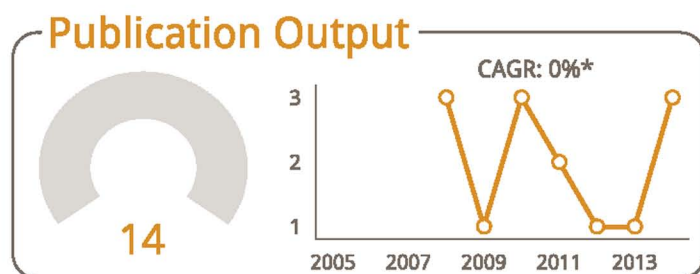


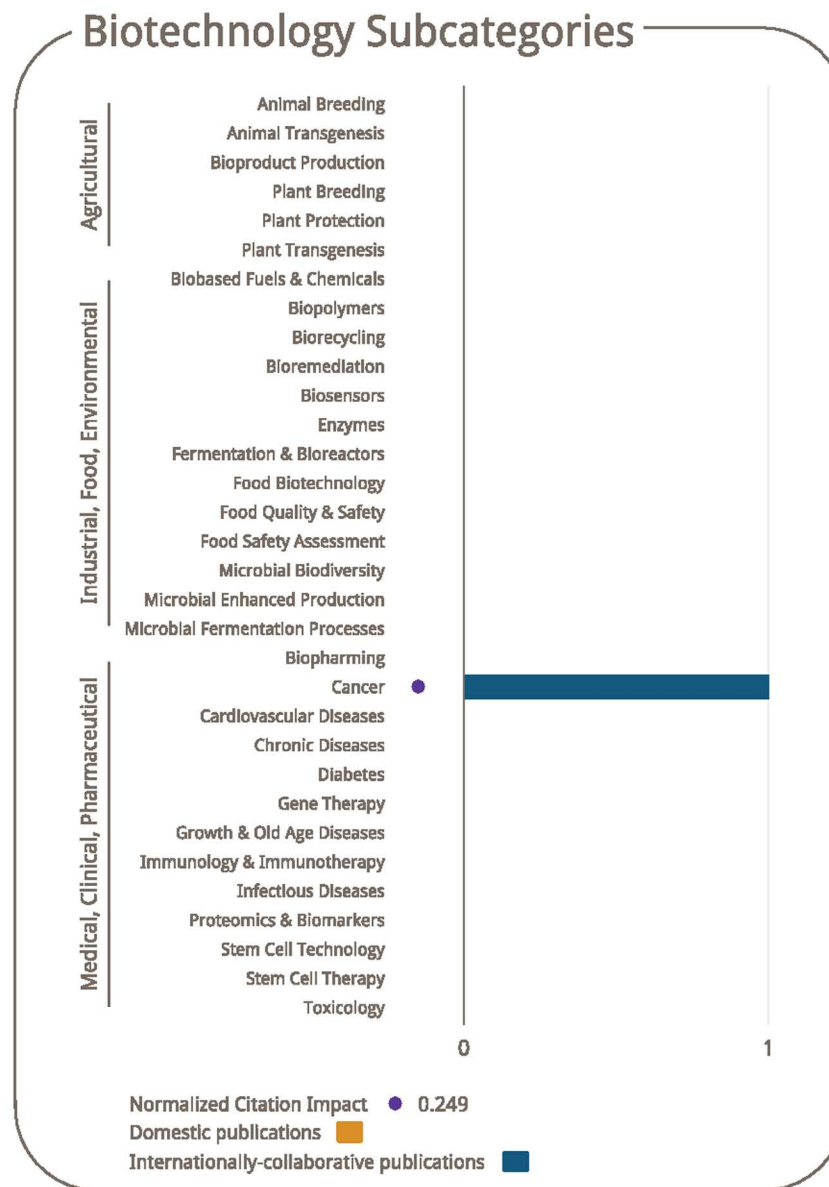
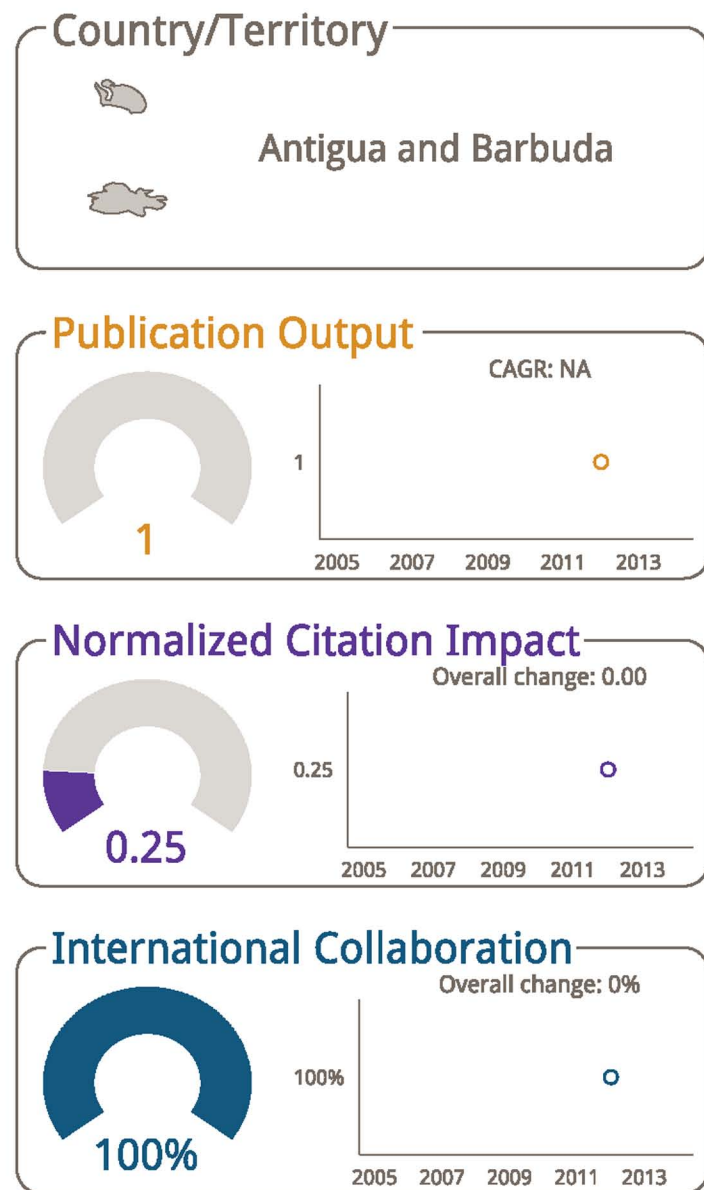
For each of the 32 **Biotechnology** subcategories, the number of domestic and internationally-collaborative publications are shown in the horizontal bar chart, and a purple circle appears next to each subcategory whose radius is proportional to the average Normalized Citation Impact of the publications in that subcategory.

A country's trends in publication output, Normalized Citation Impact (NCI) and International Collaborations are shown in the linechart. The growth in publication output is measured by the compound annual growth rate (CAGR, calculated by using  $\frac{\text{End year value}}{\text{Start year value}}^{\frac{1}{\# \text{ of years}}} - 1$ ), and the **Overall Change** in NCI and international collation rate is calculated by directly subtracting the start year value from the end year value. The start year was the earliest year of publication, and the end year was the most recent publication year up to 2014. The CAGR was not calculated (shown as "NA") for those countries which had sporadic (discontinuous) publications. The asterisked CAGRs refer to the cases when the continuous activity was observed in only one time interval but shorter than 2005-2014.

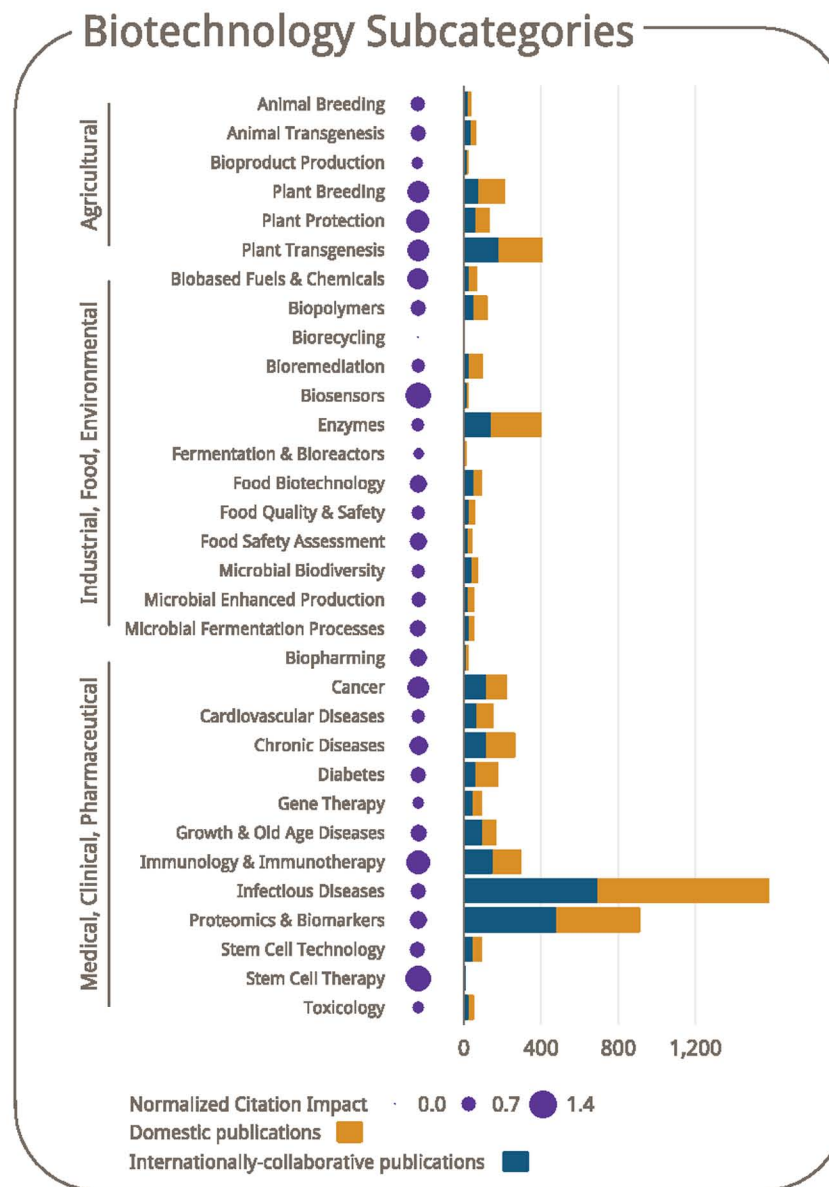
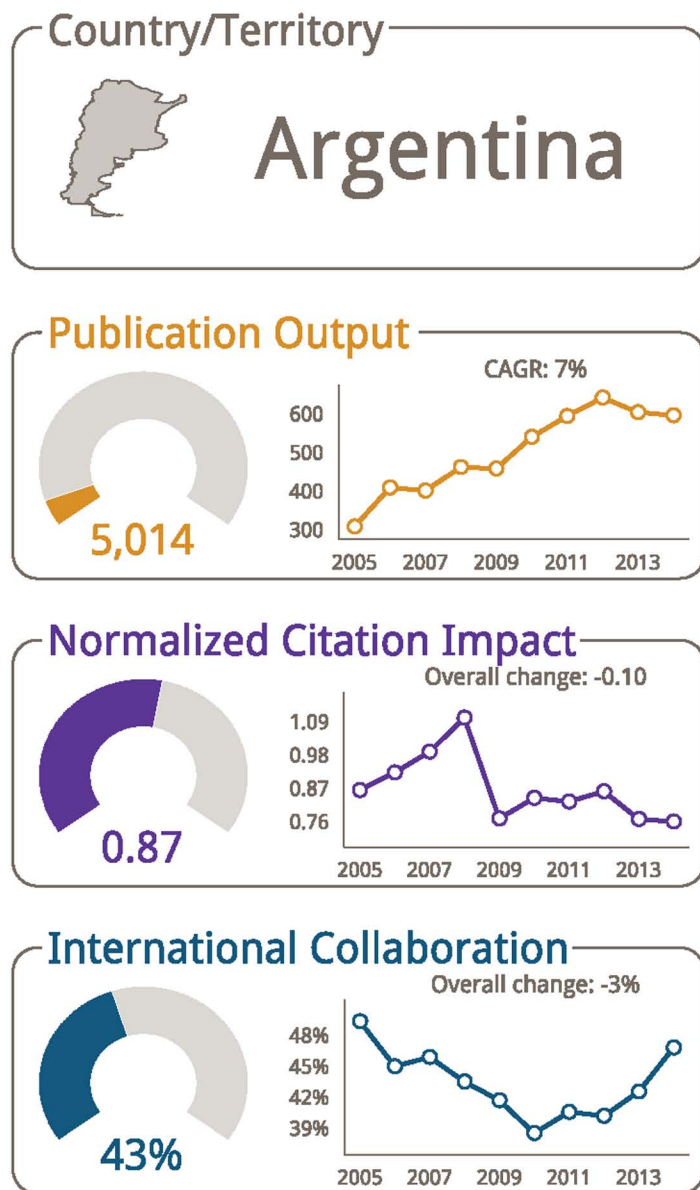


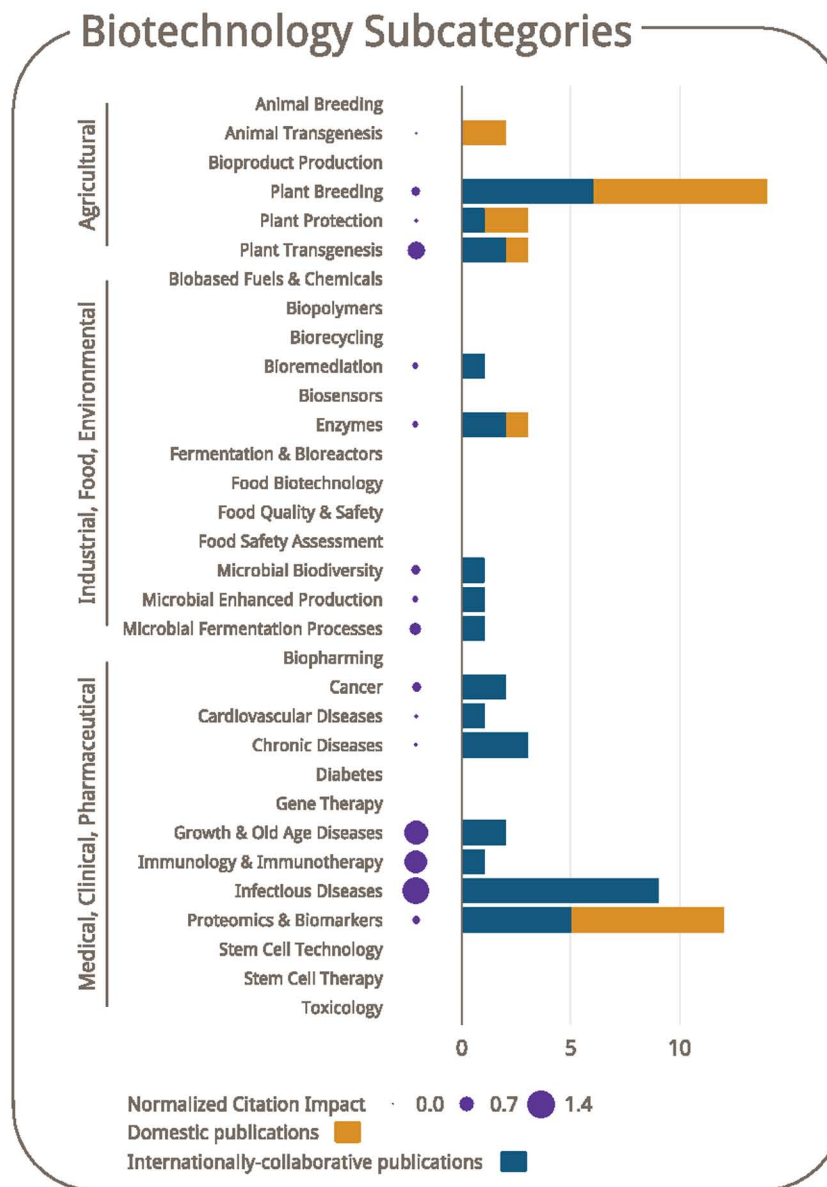
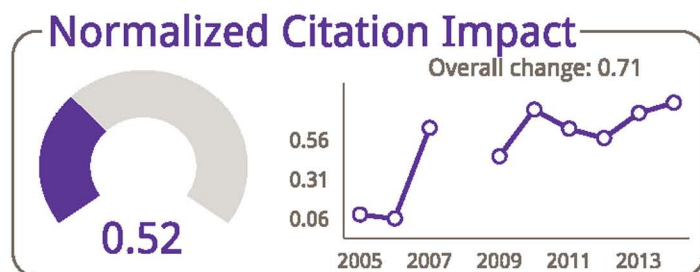
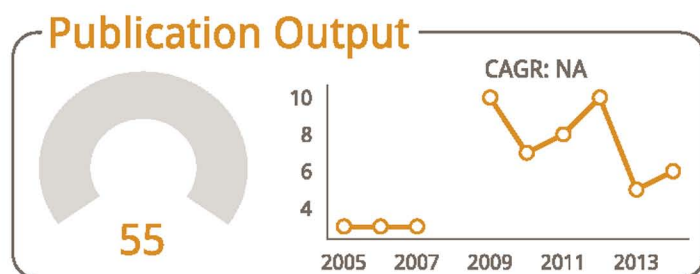


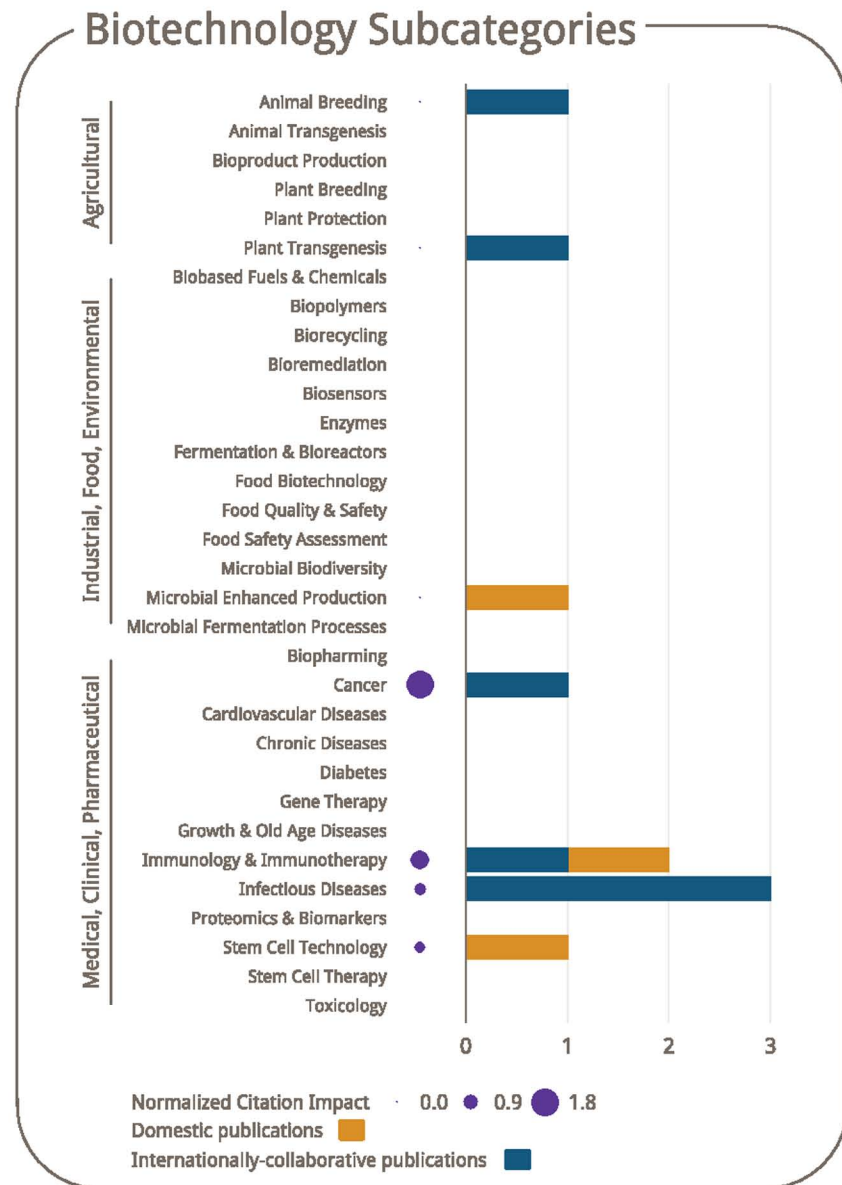
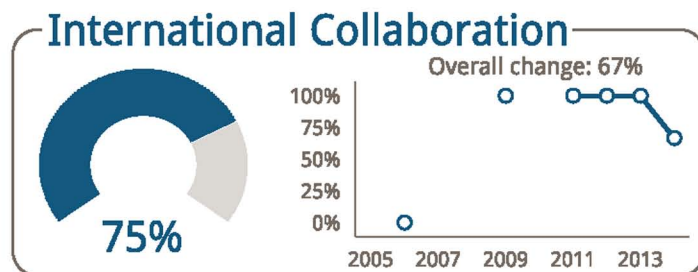
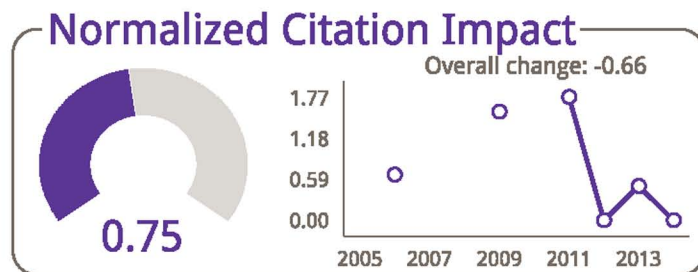
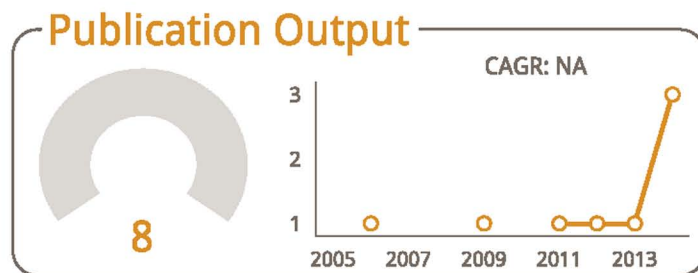




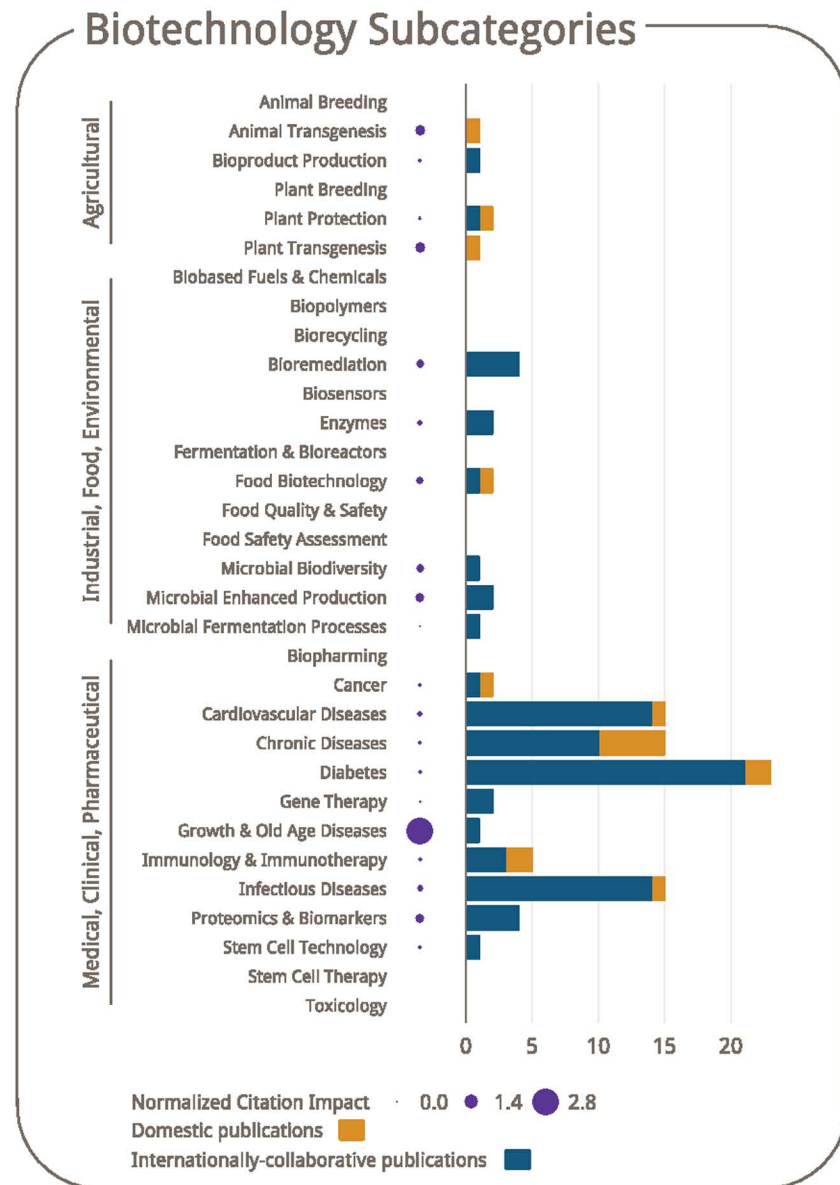
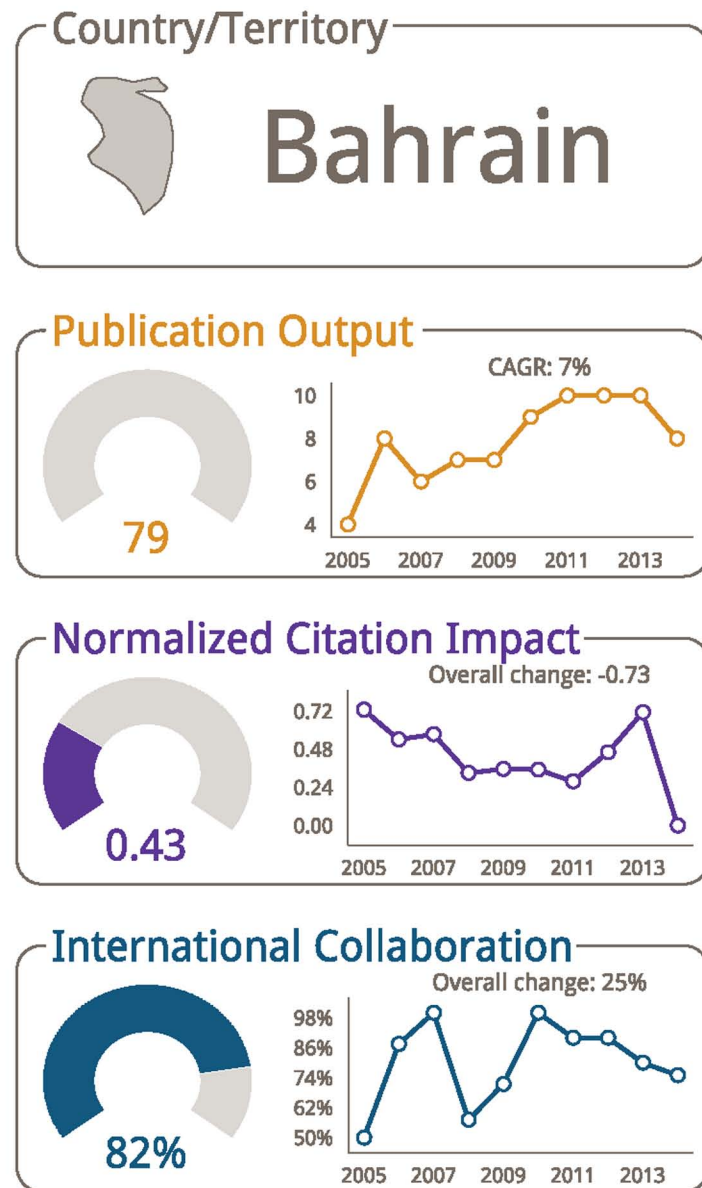


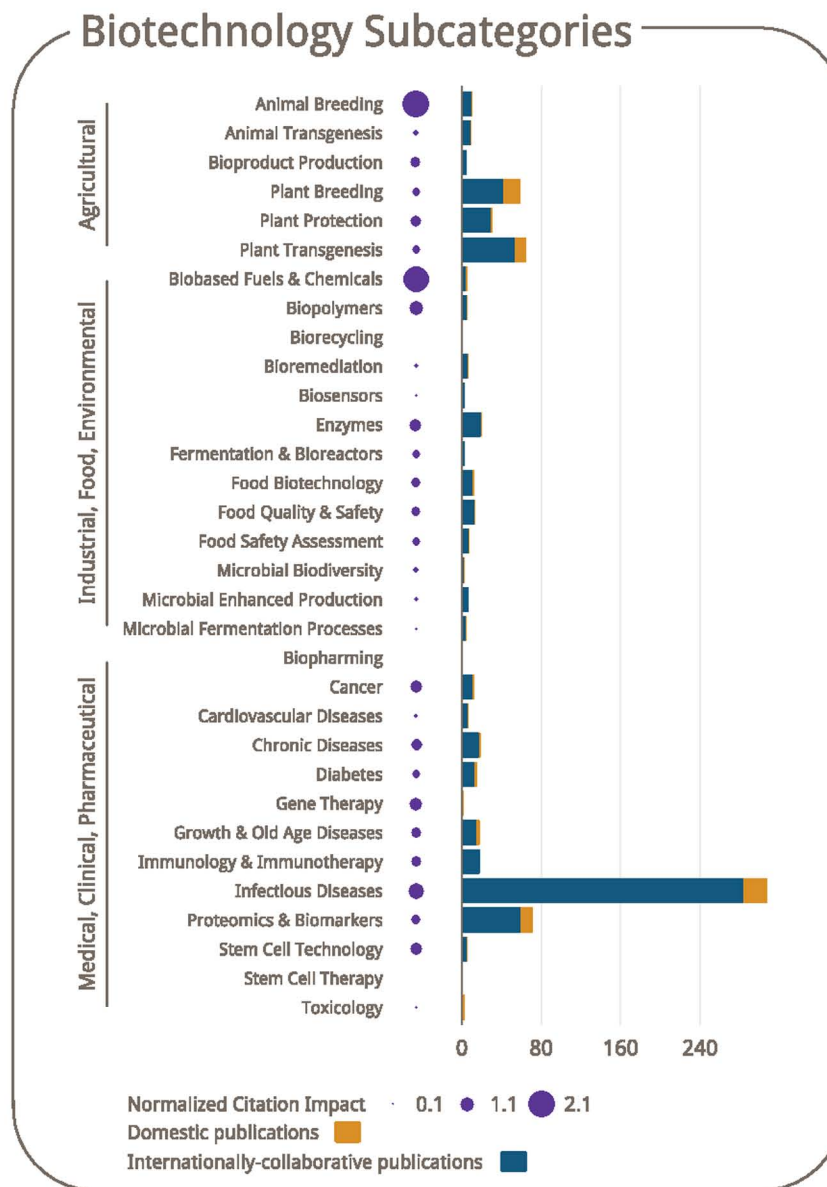
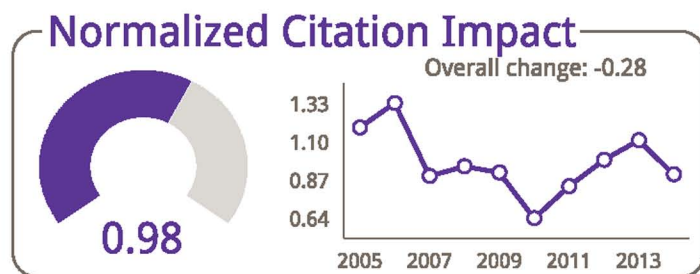
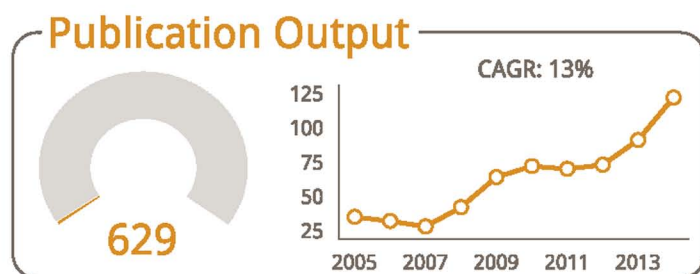


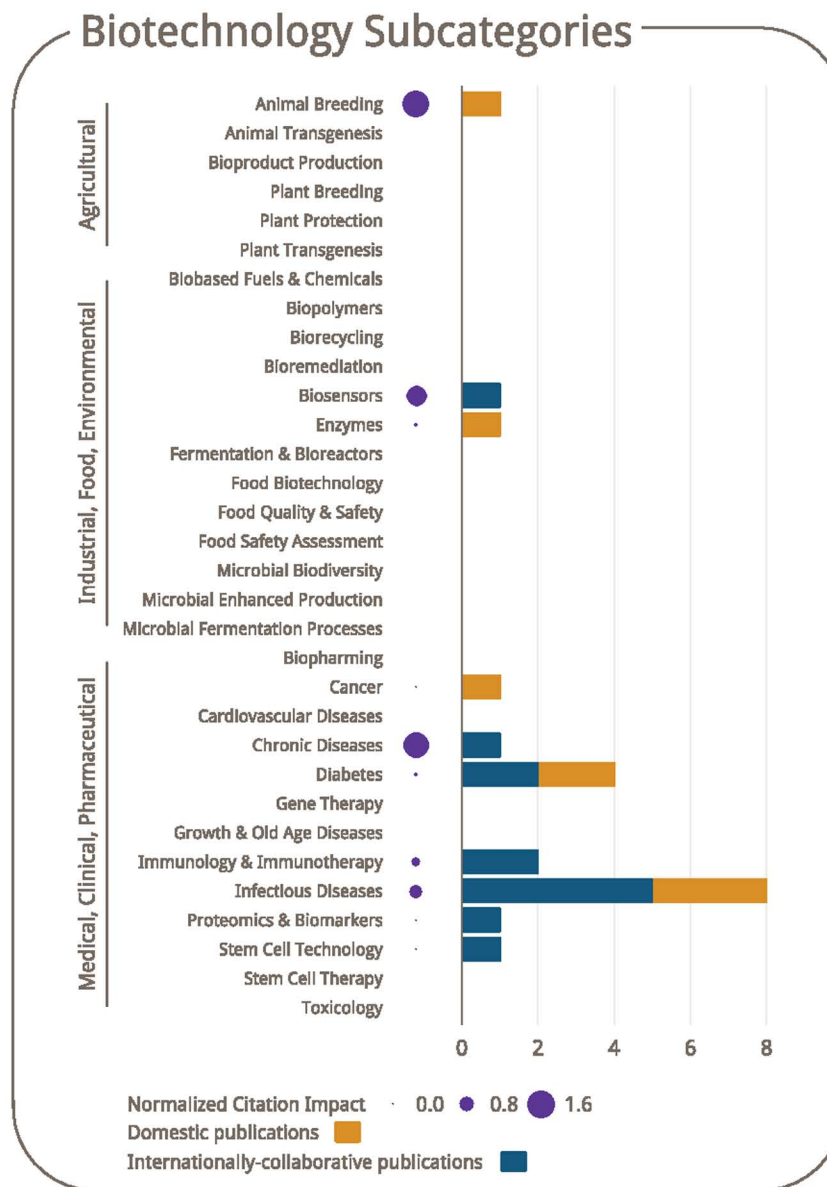
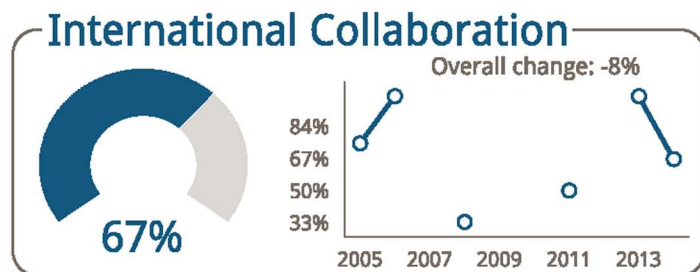
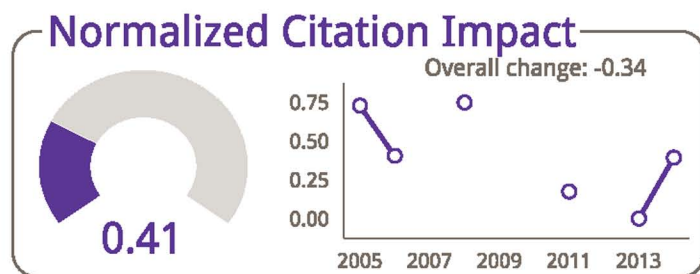
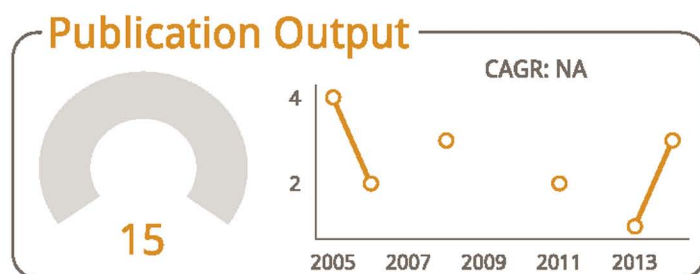


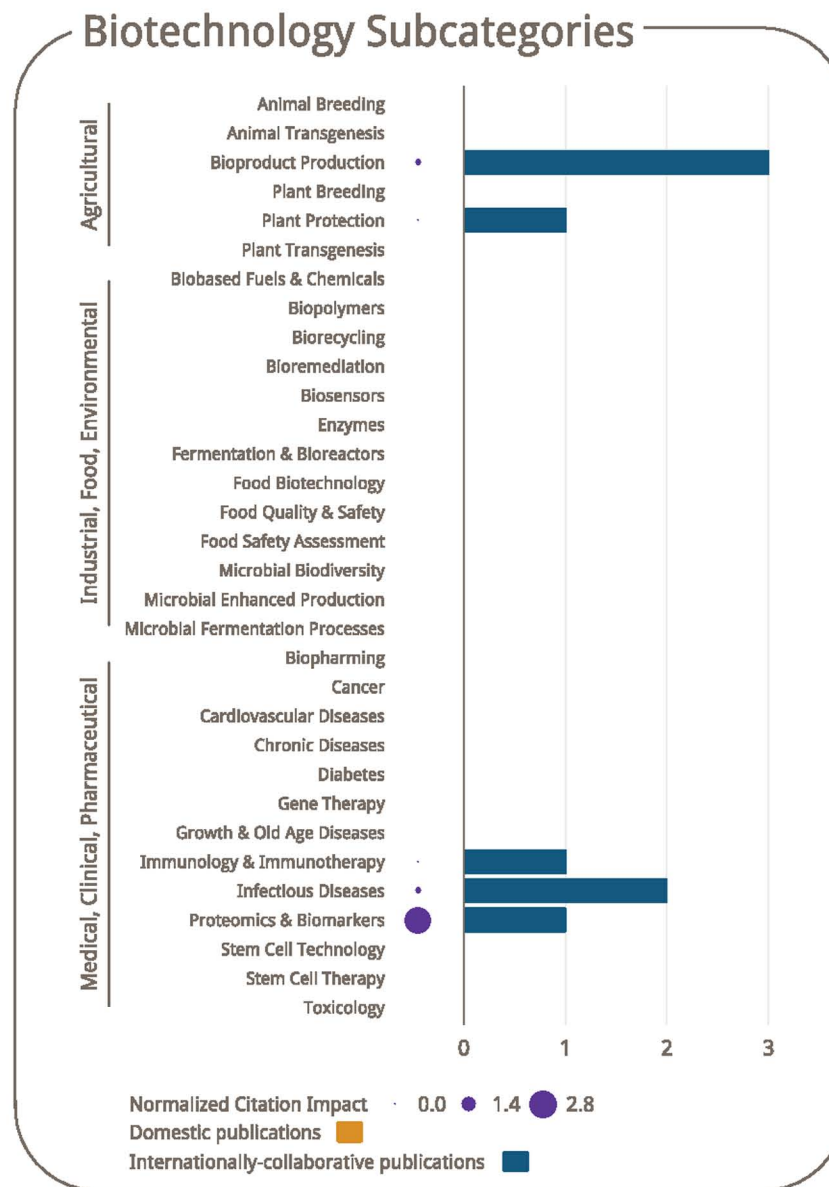
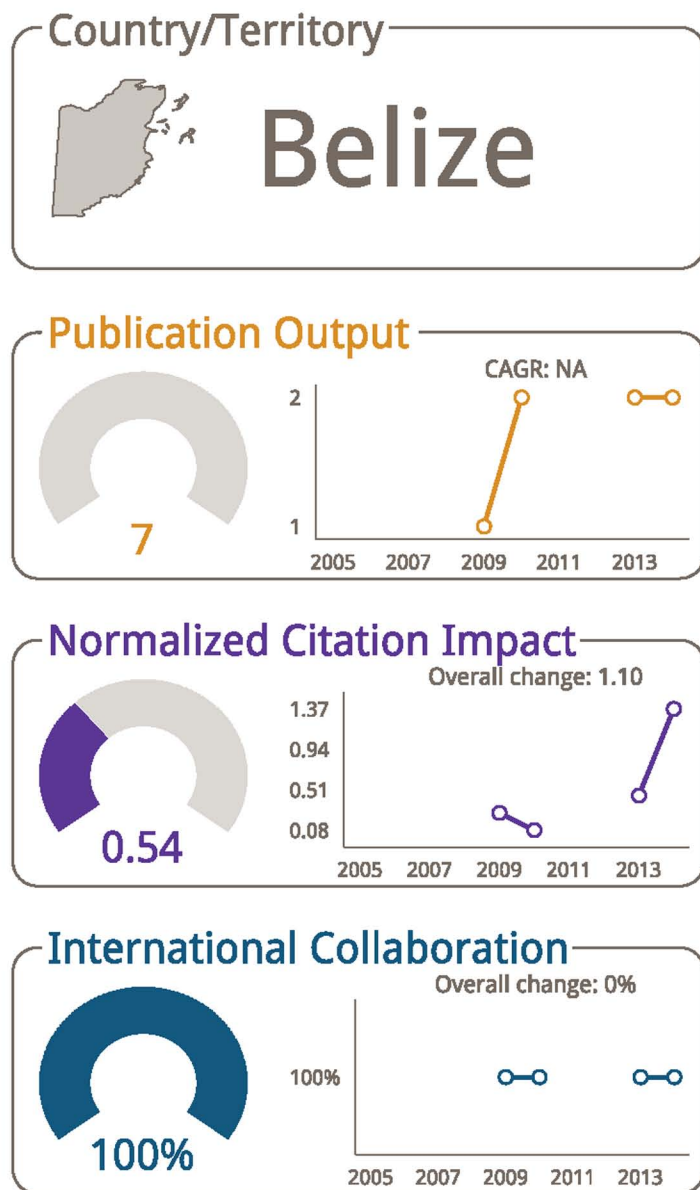


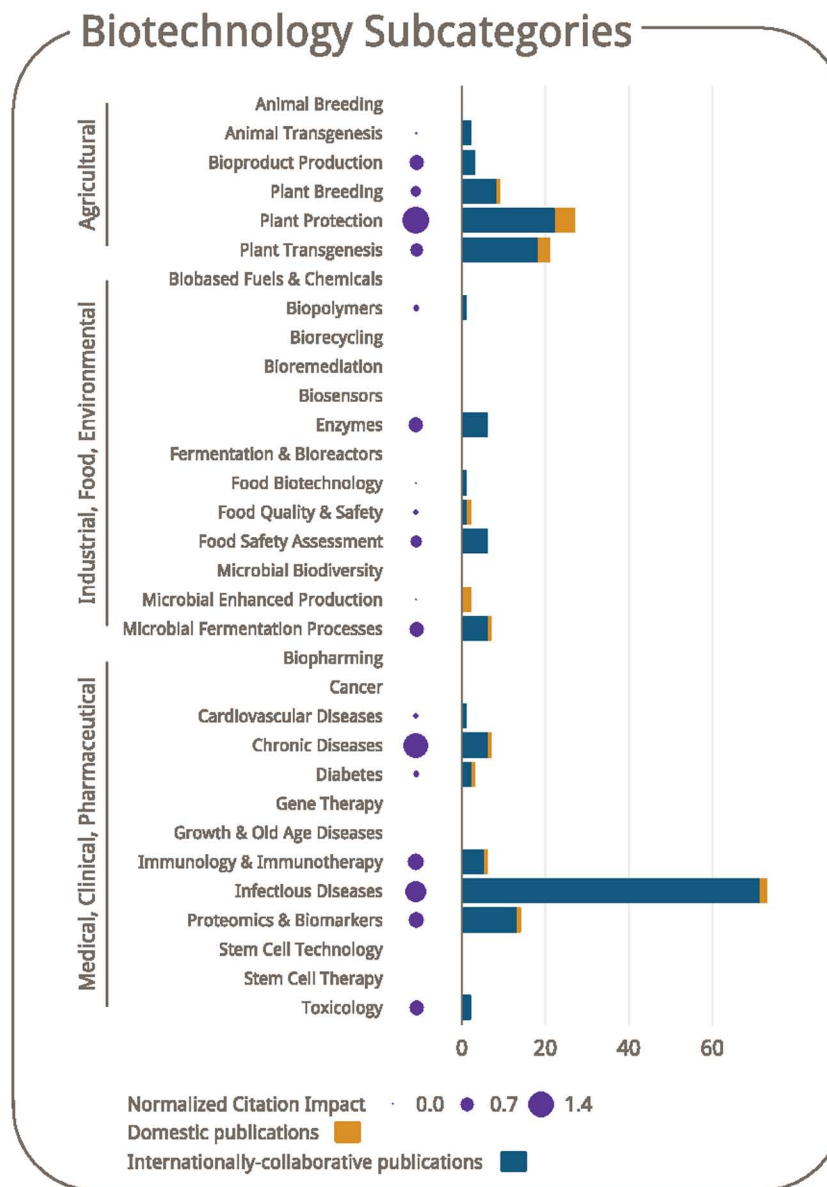
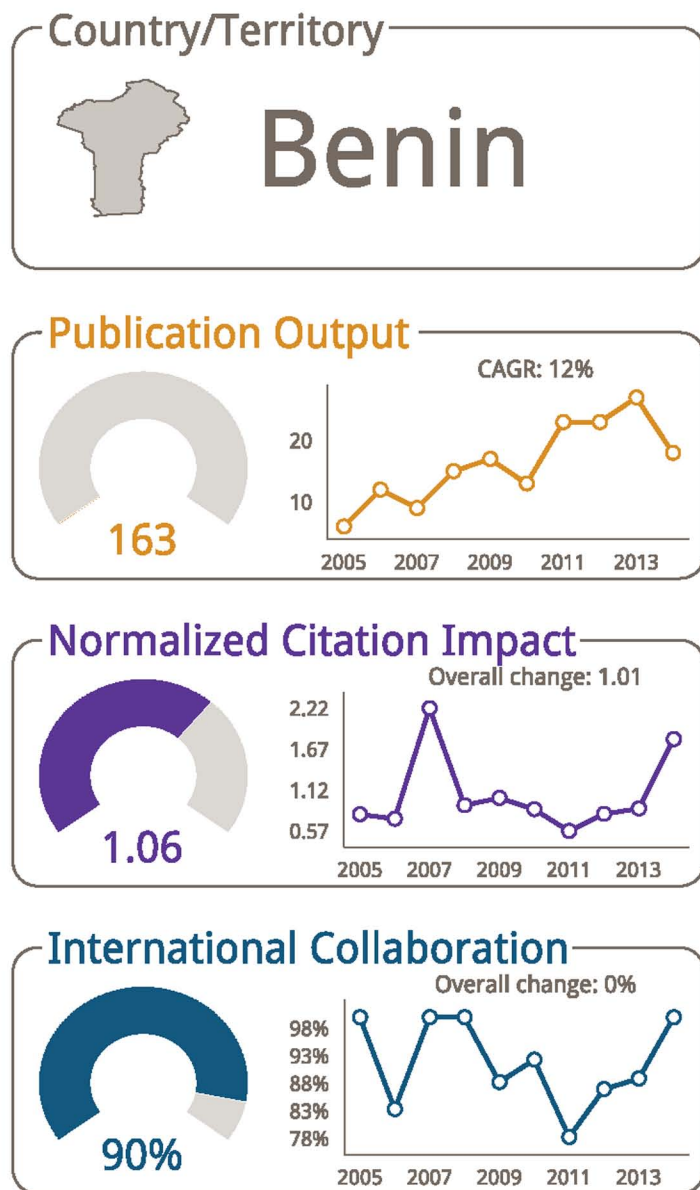




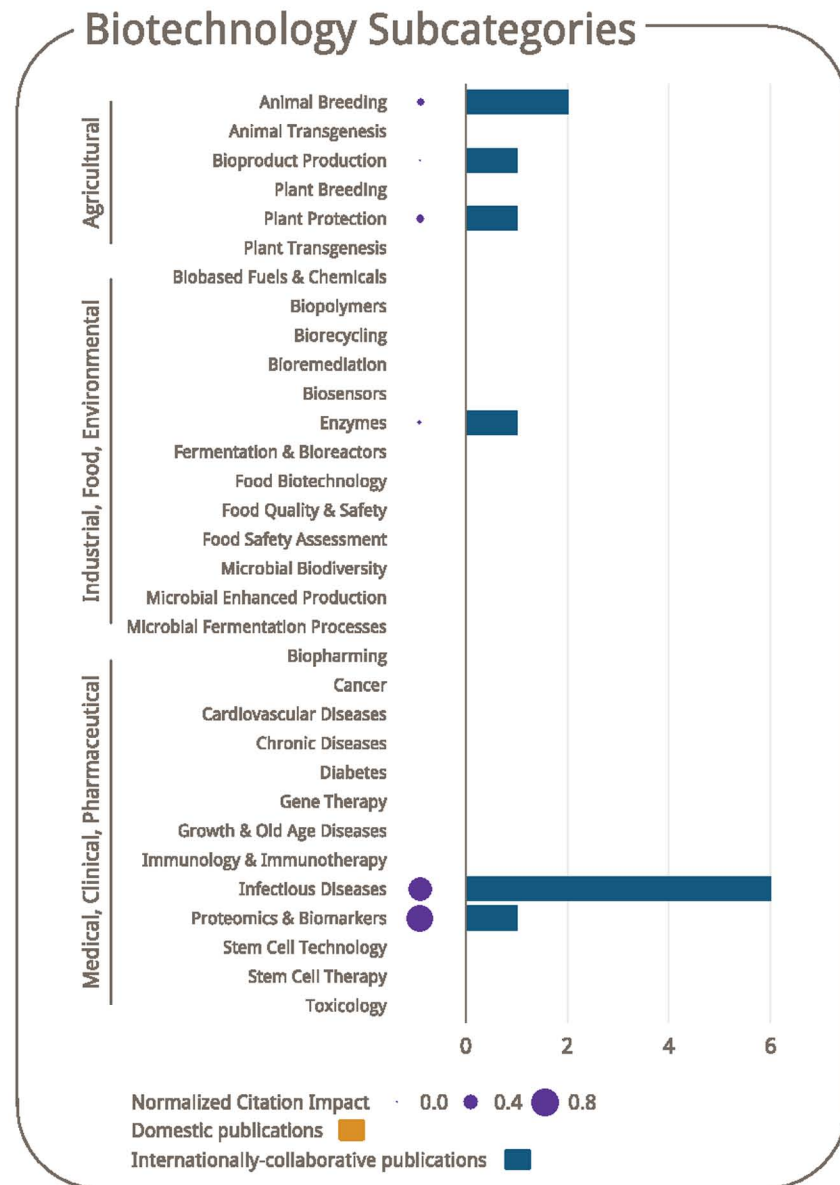
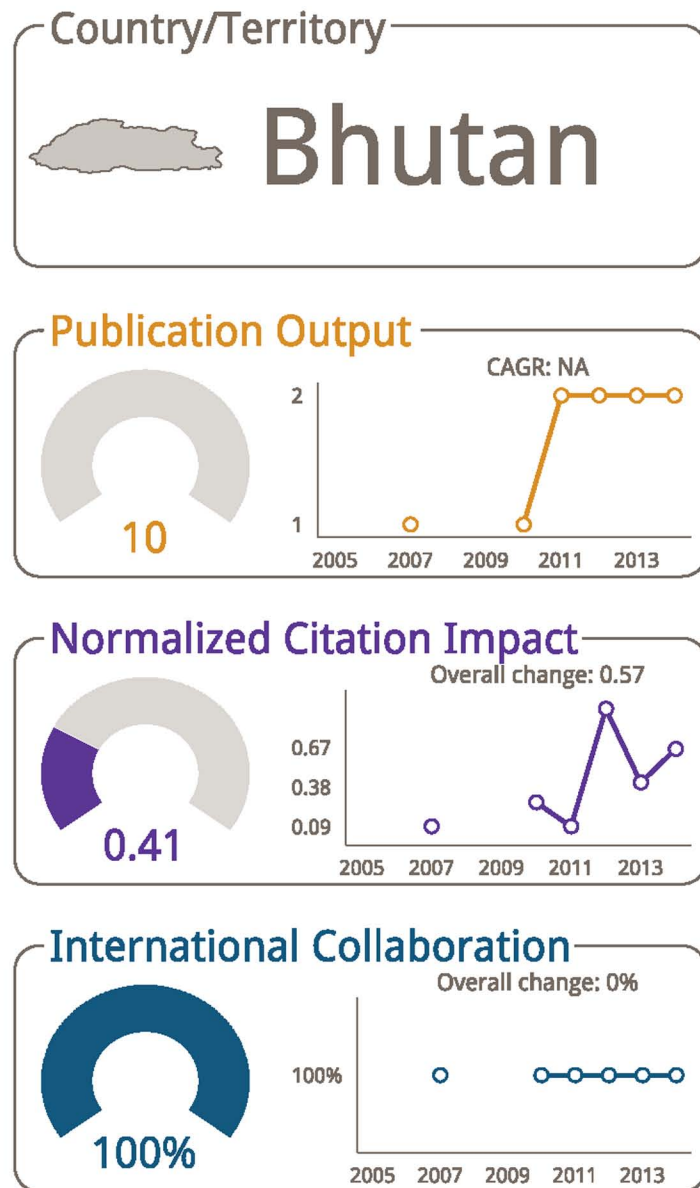


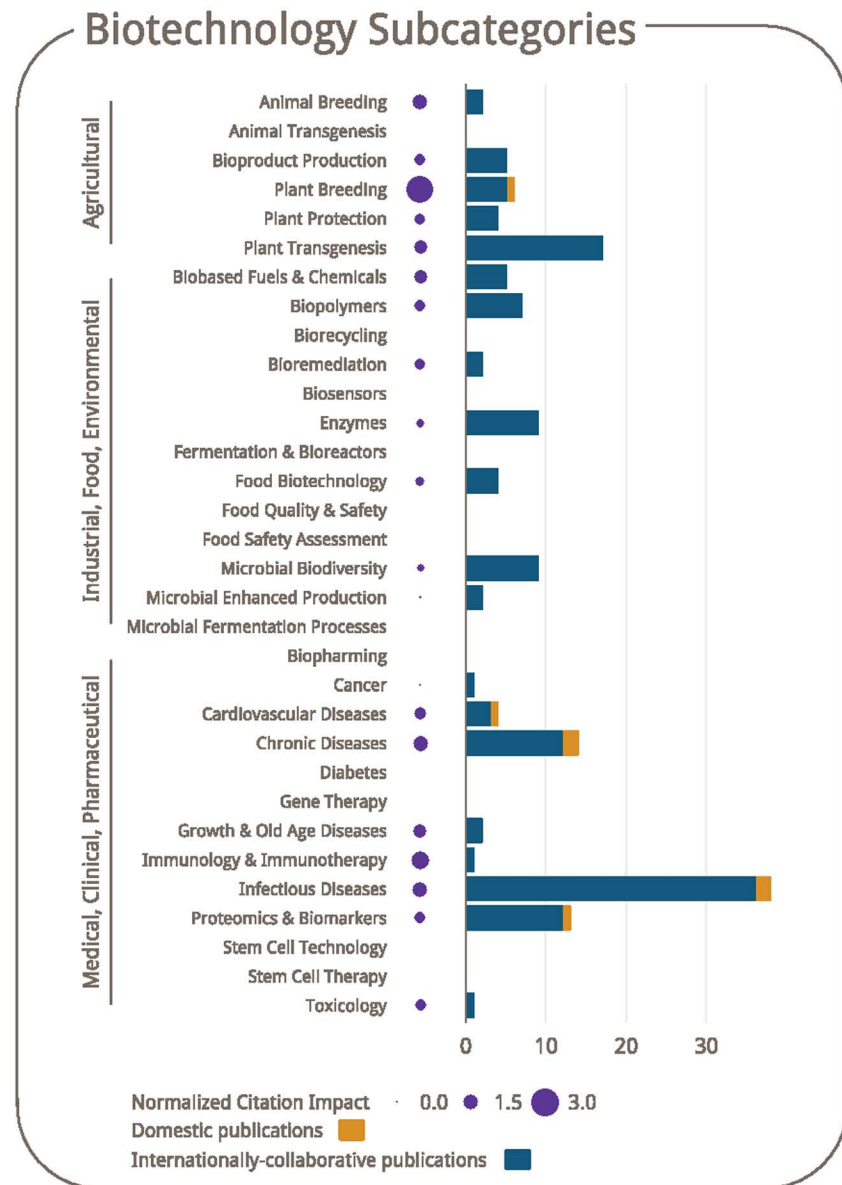
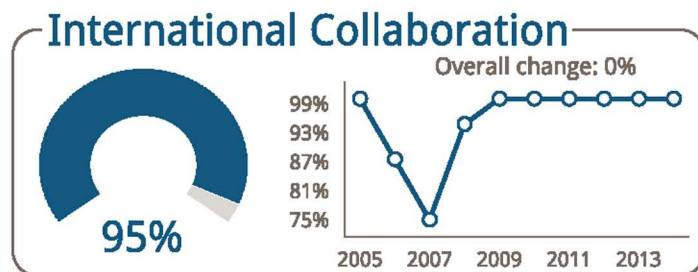
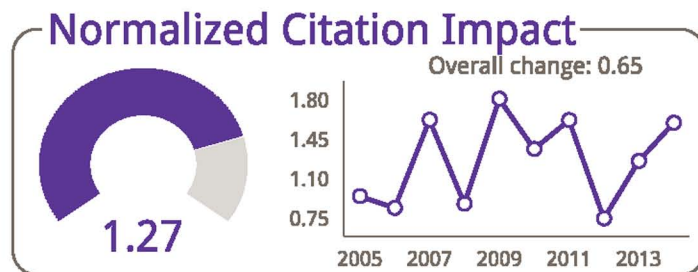
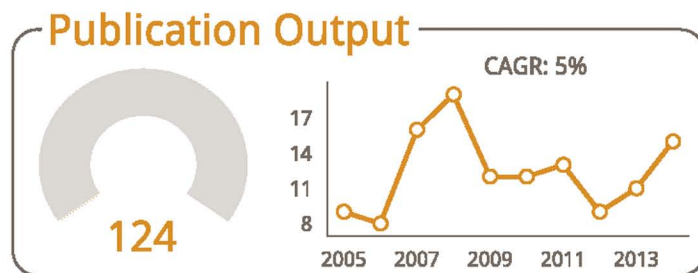


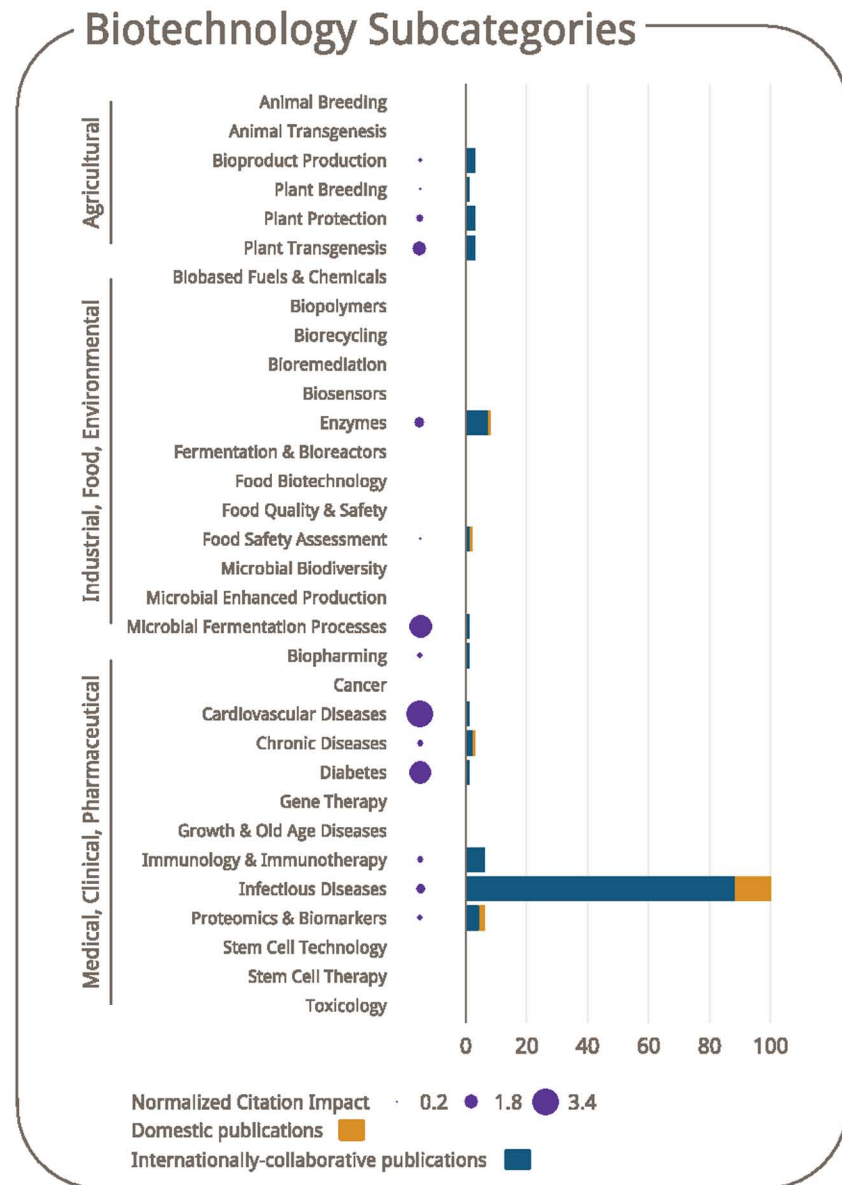
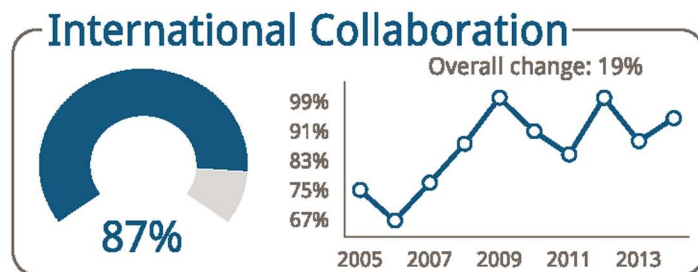
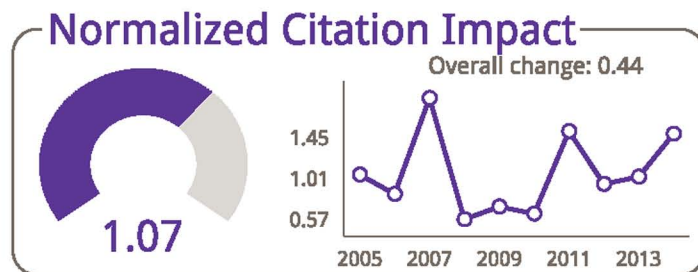
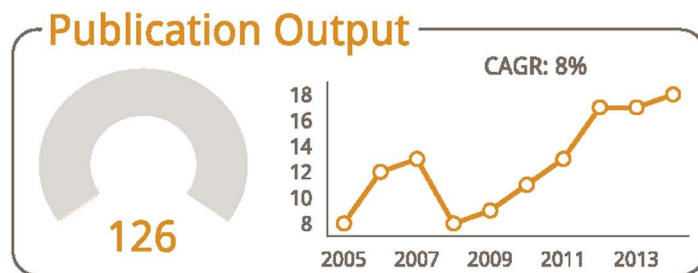




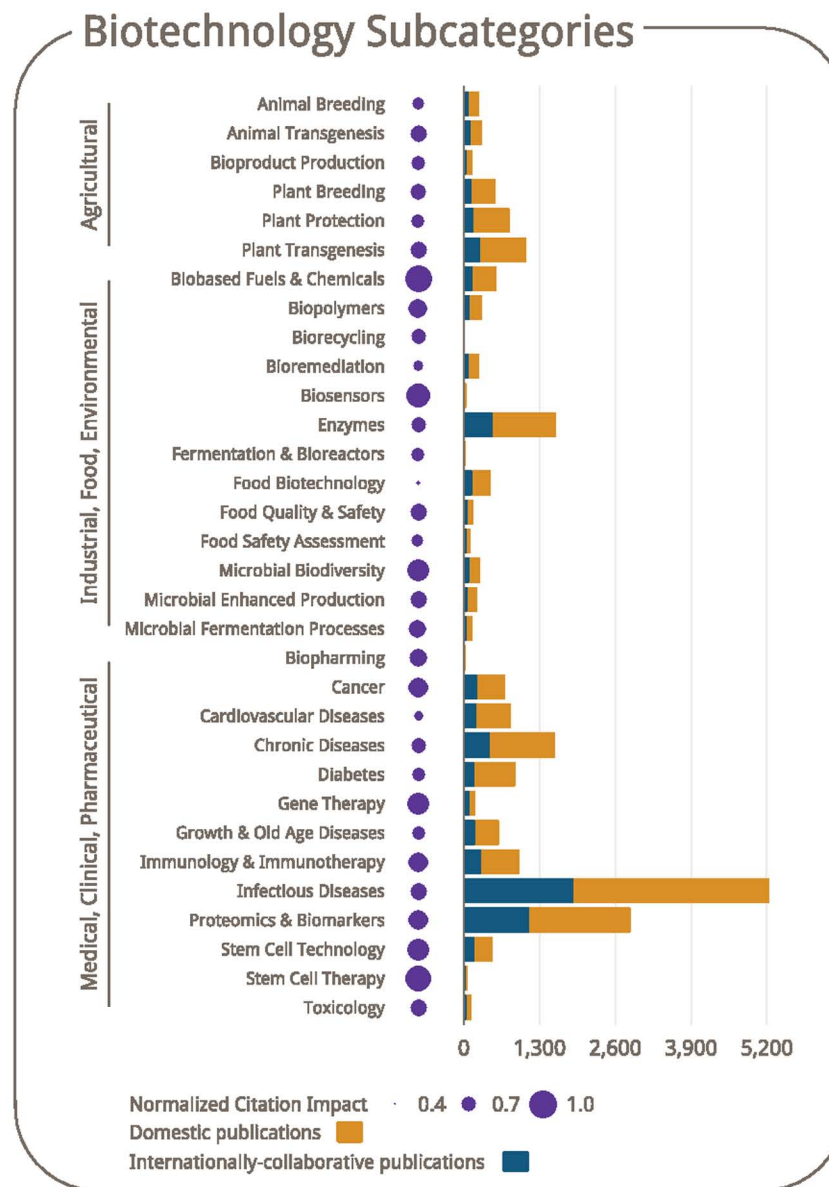
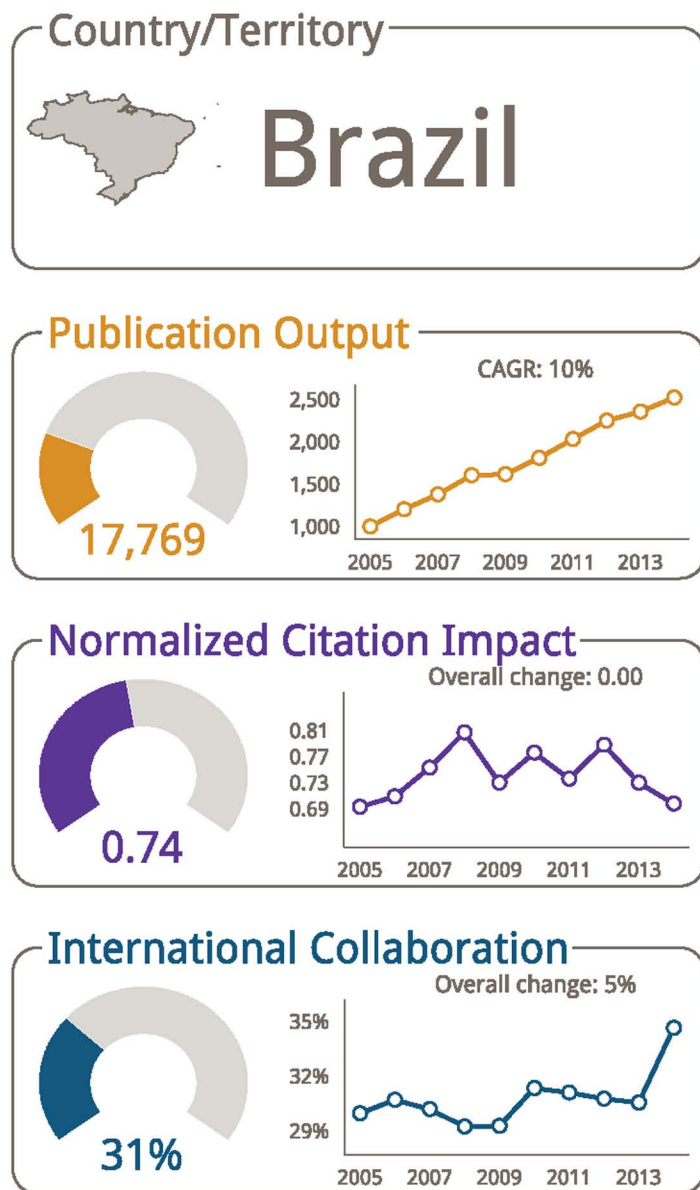










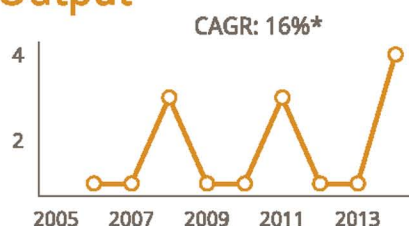


## Country/Territory

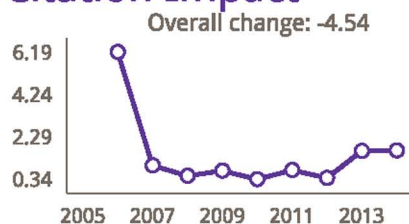


Brunei Darussalam

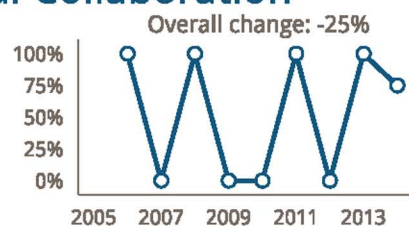
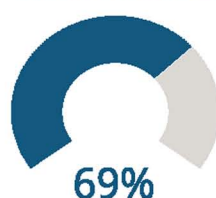
## Publication Output



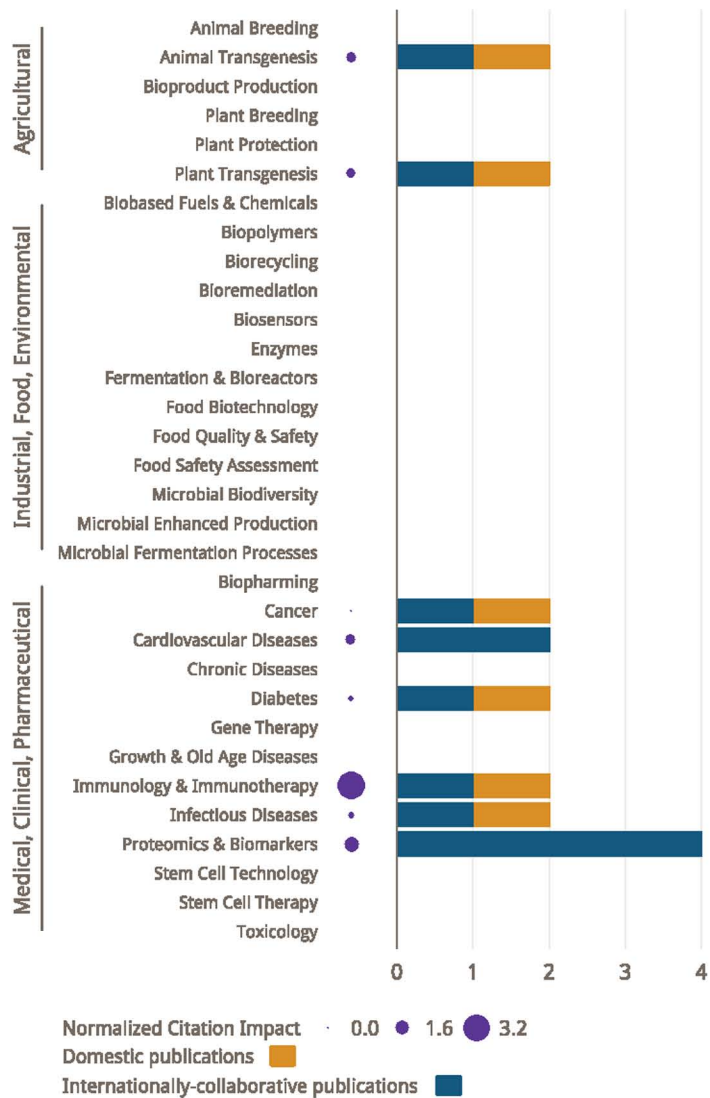
## Normalized Citation Impact

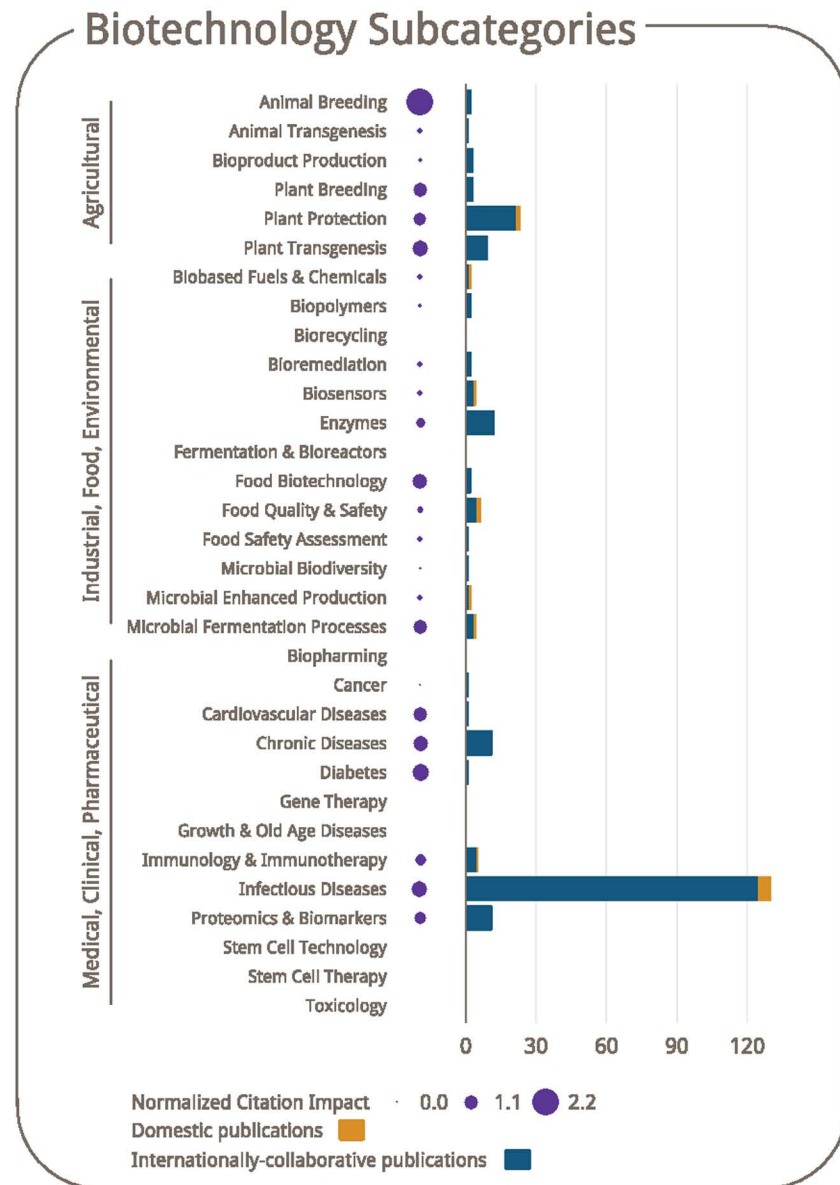
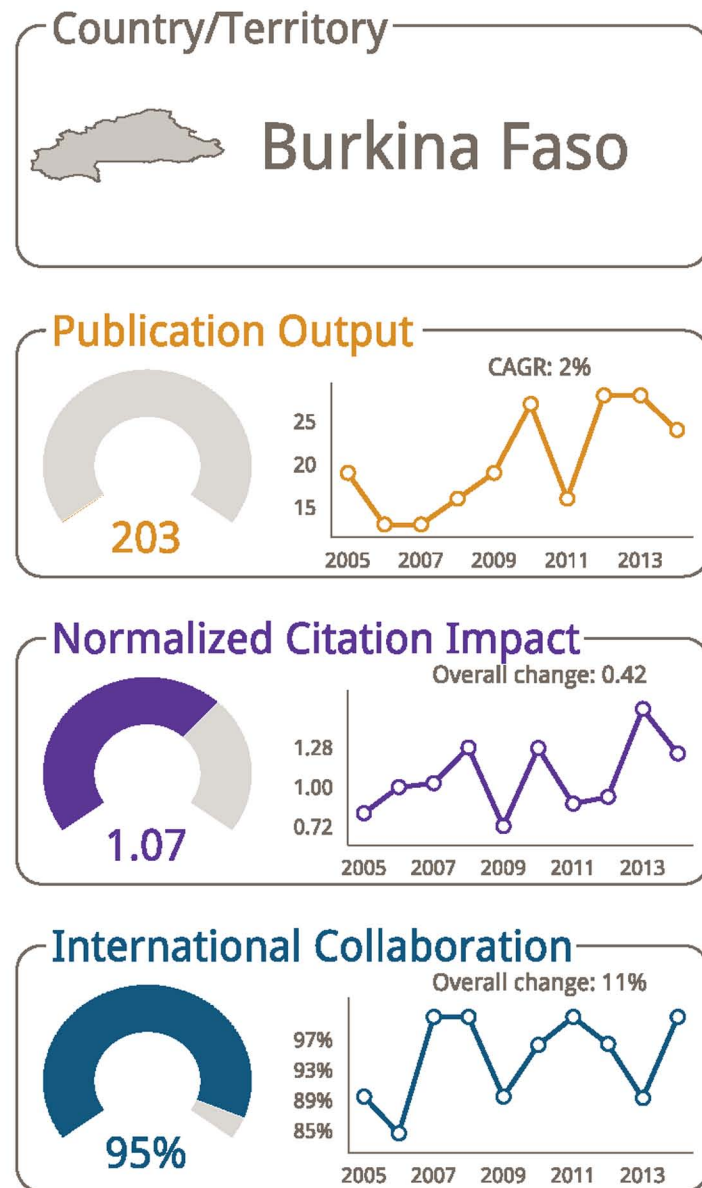


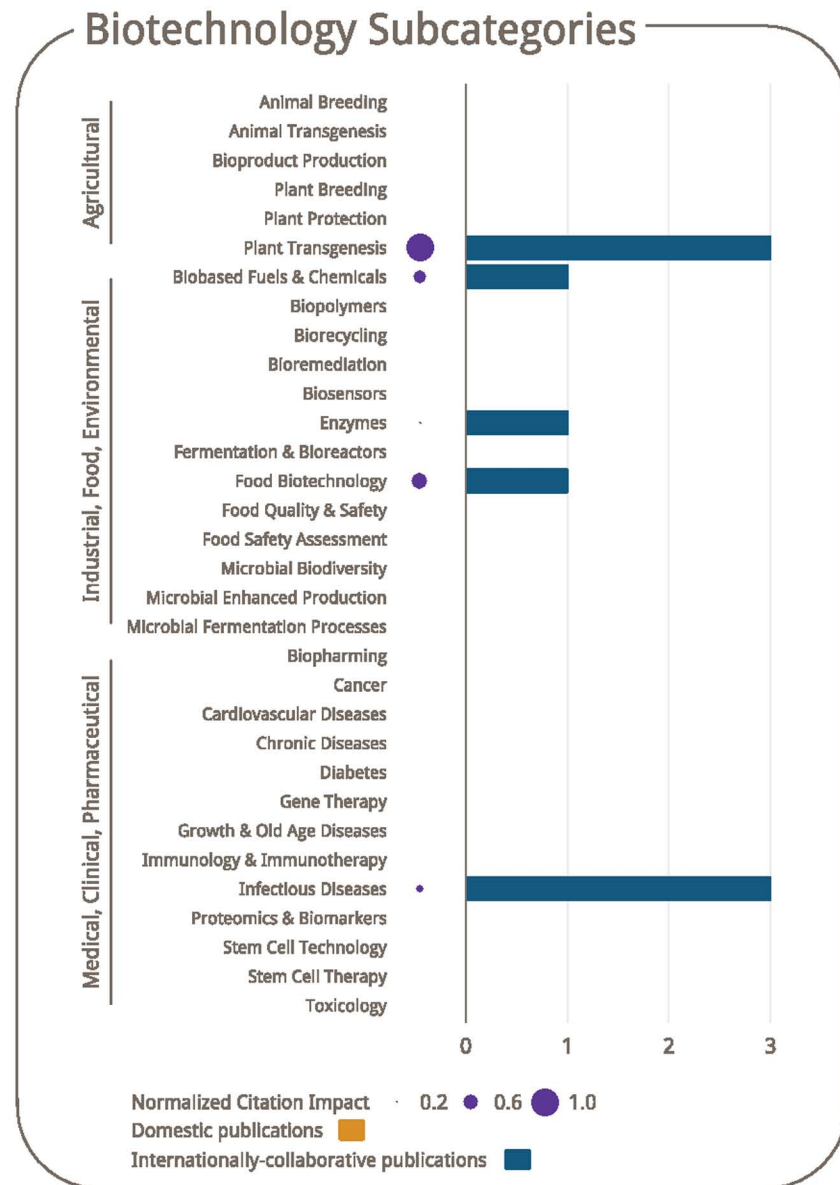
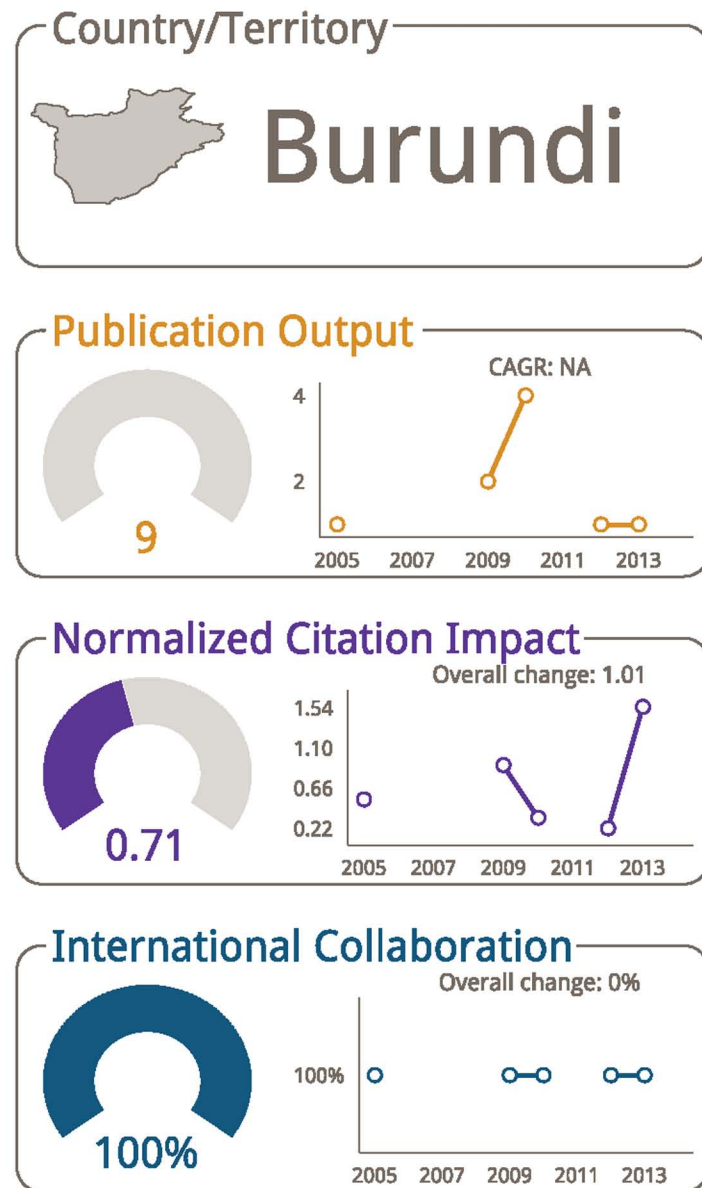
## International Collaboration

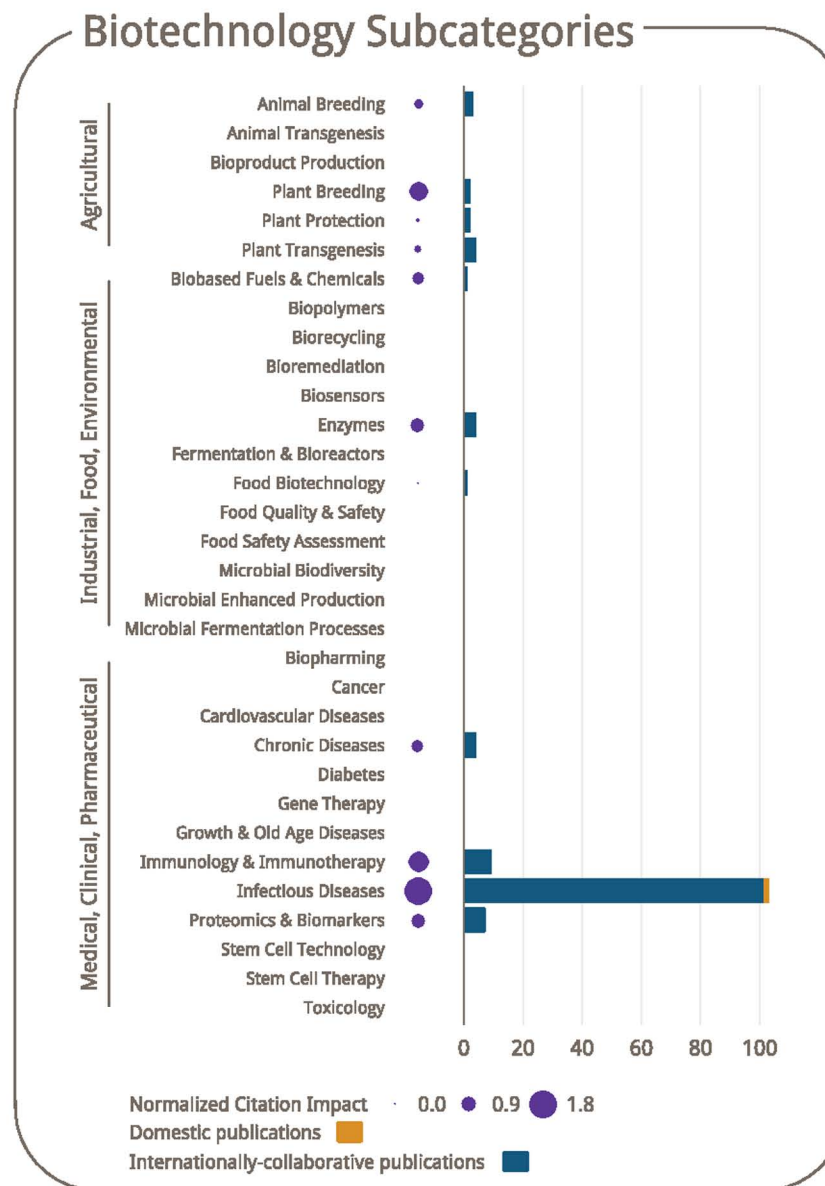
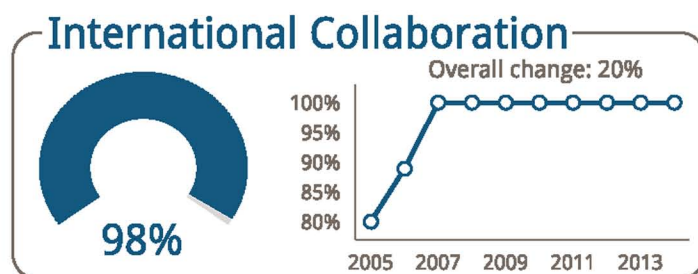
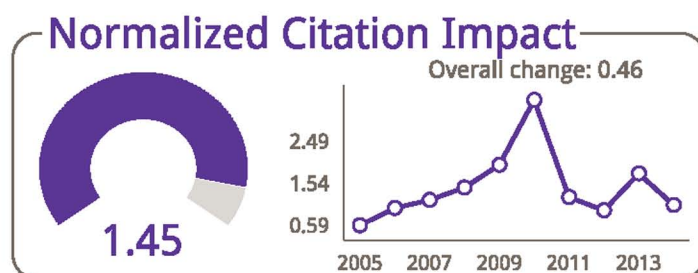
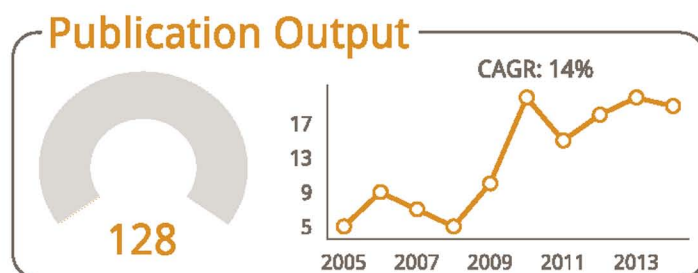


## Biotechnology Subcategories

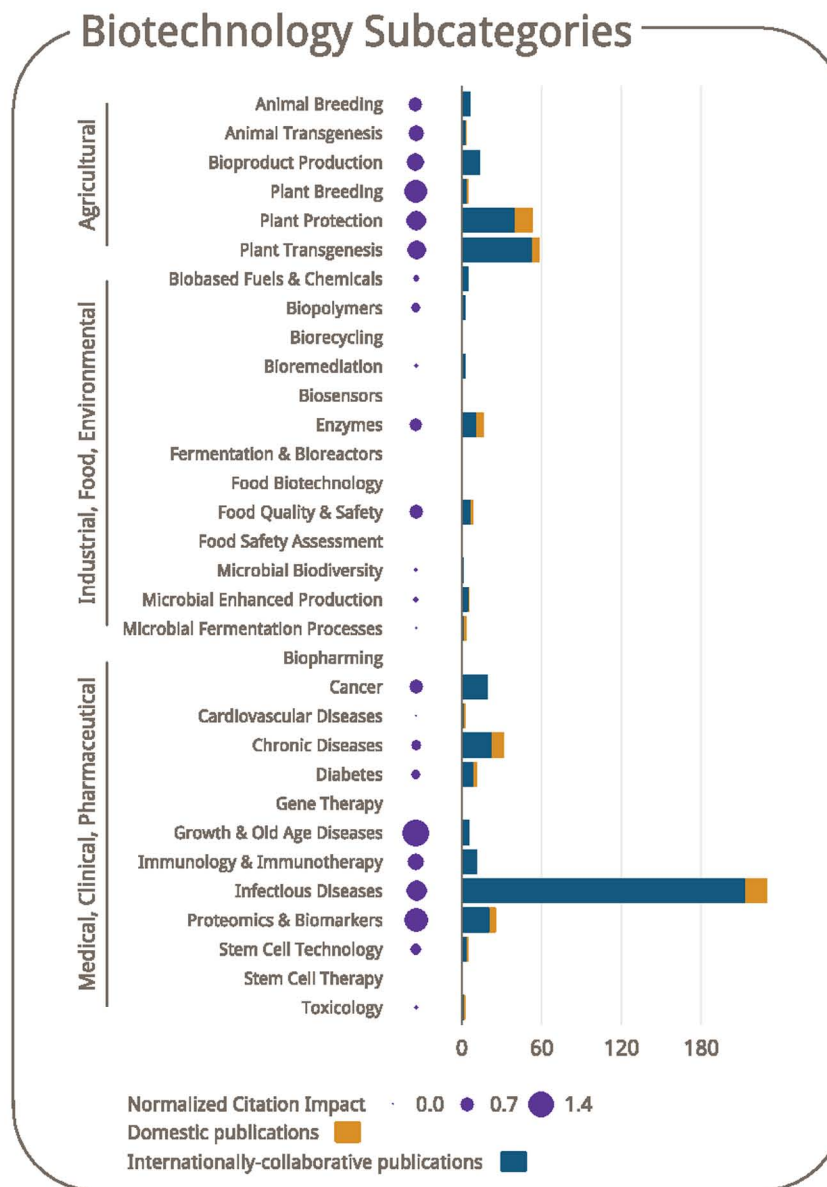
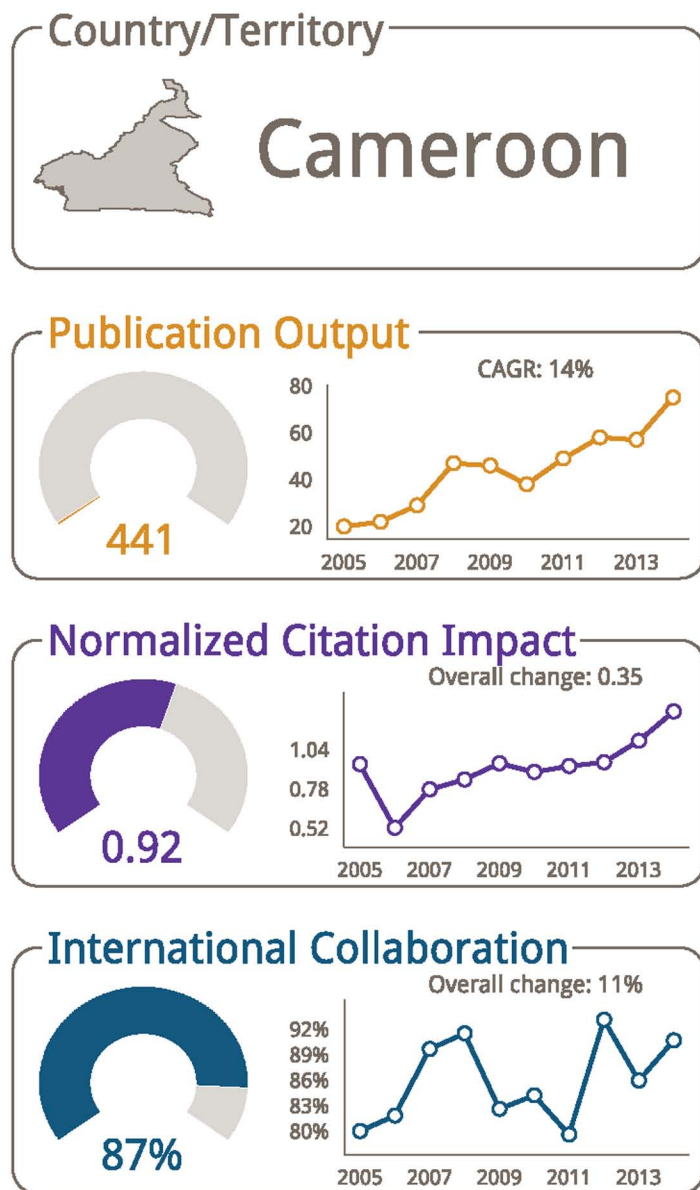


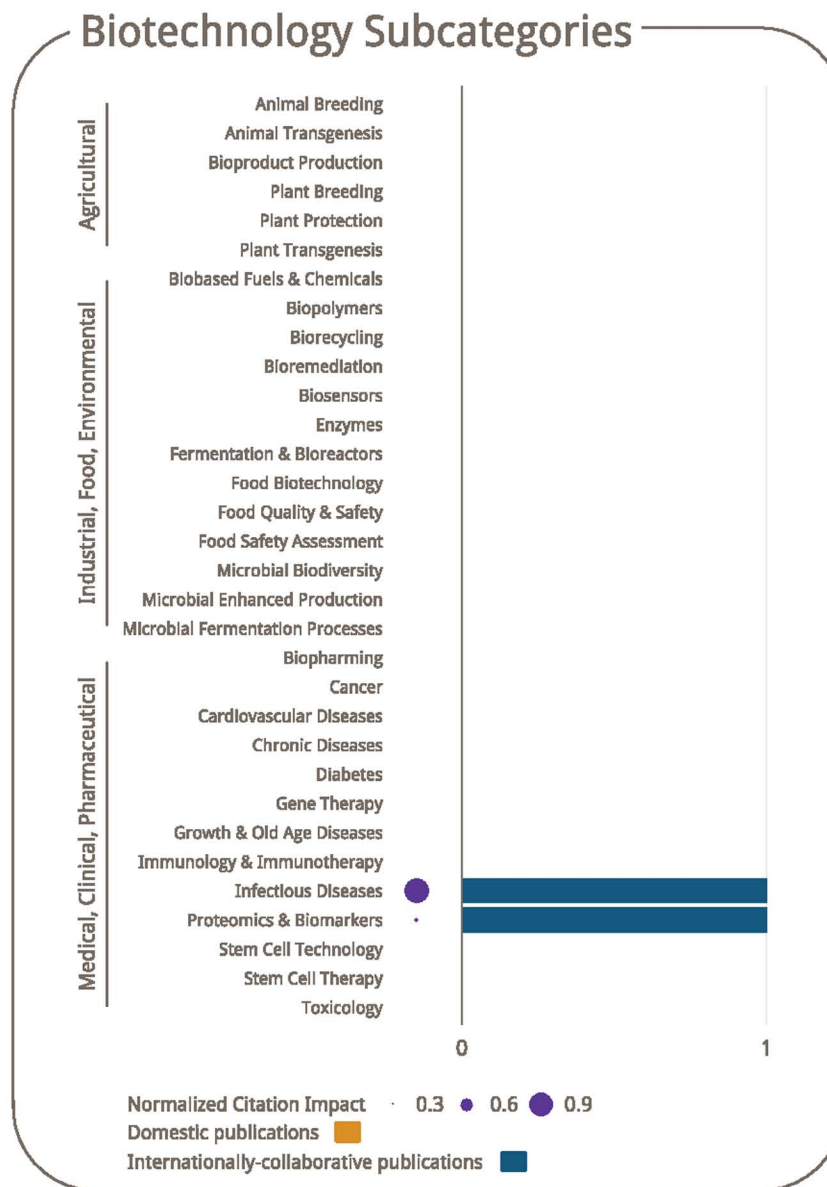
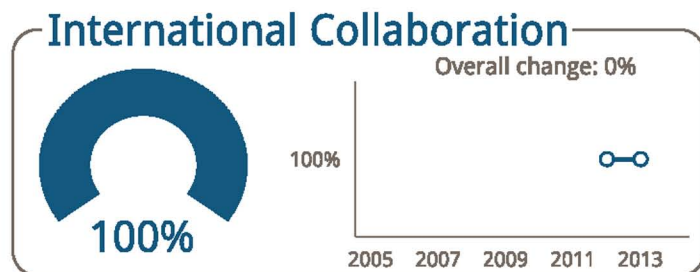
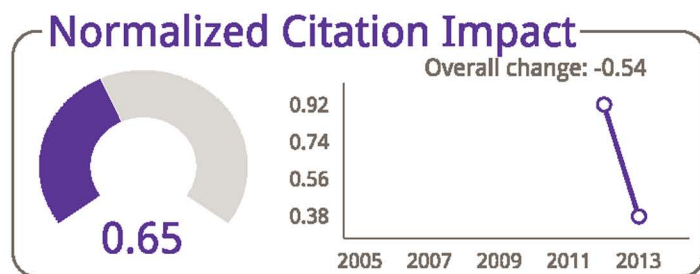
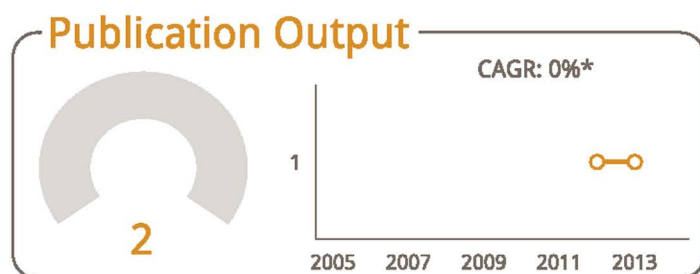


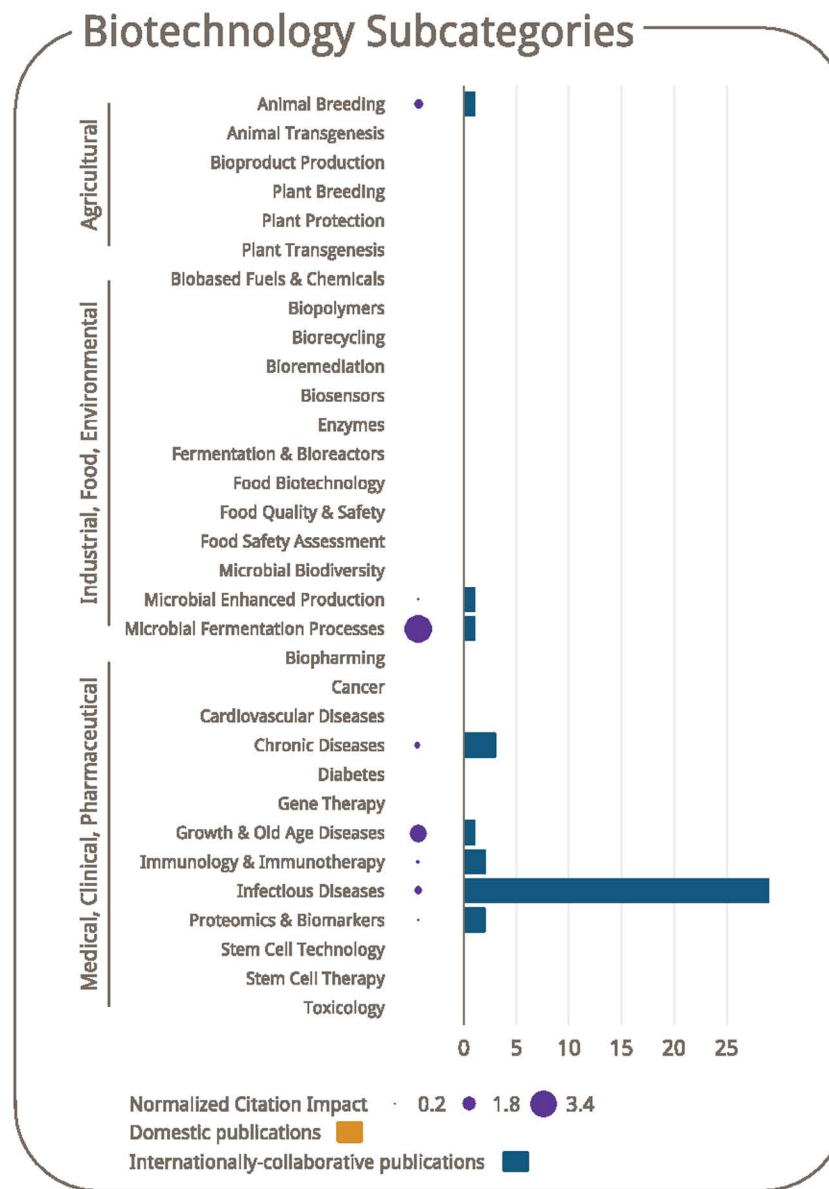
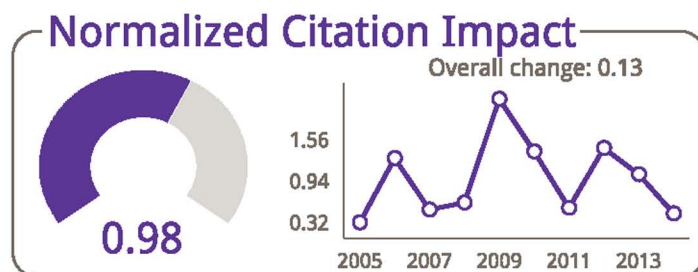
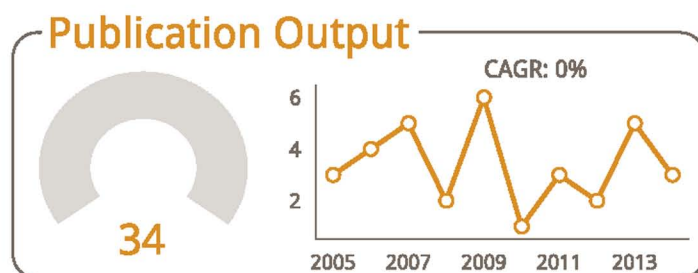




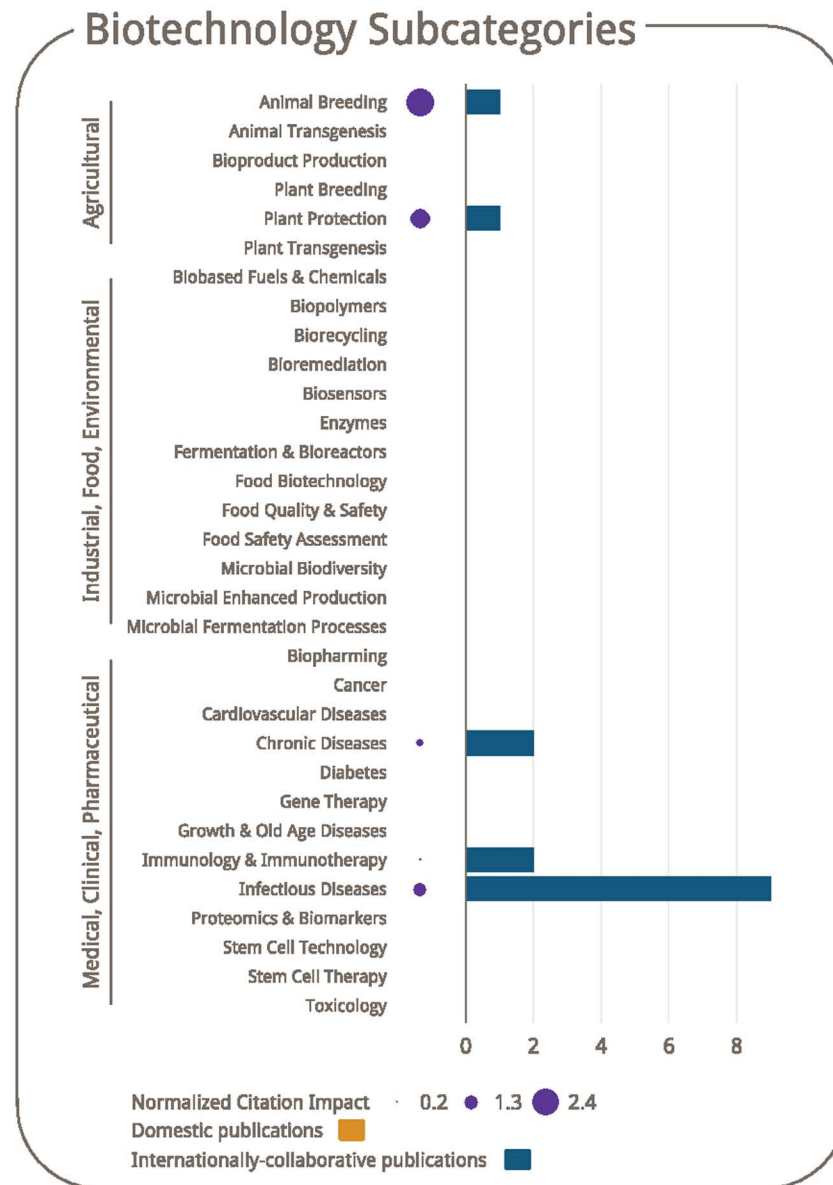
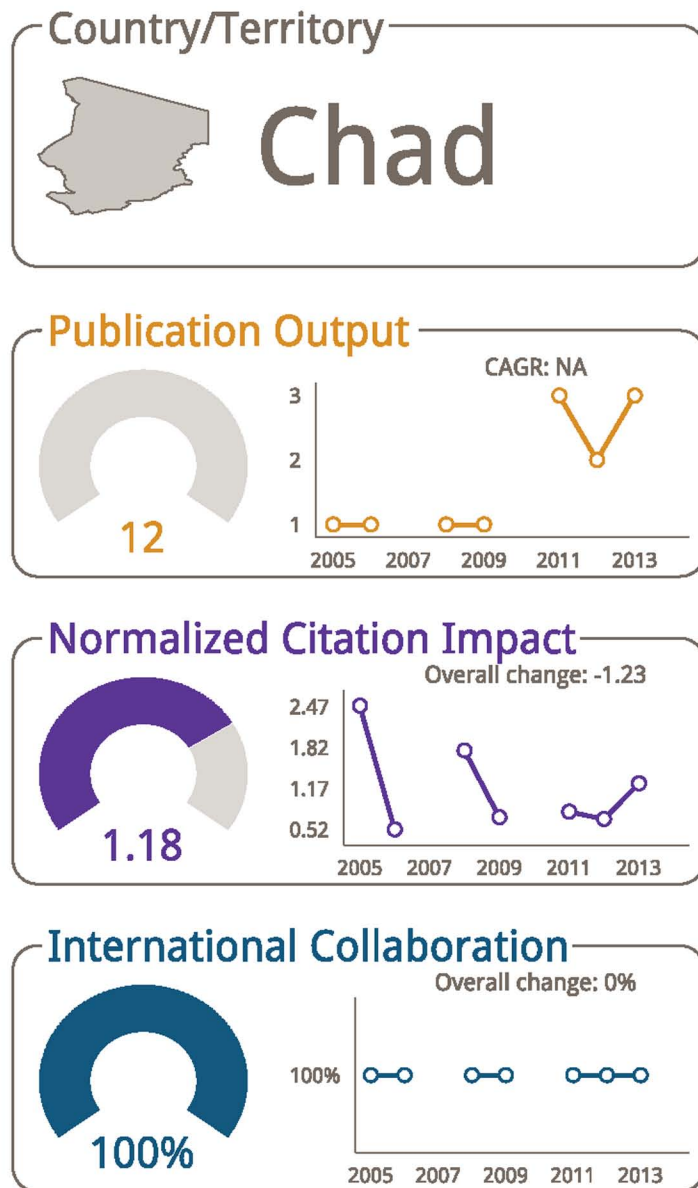


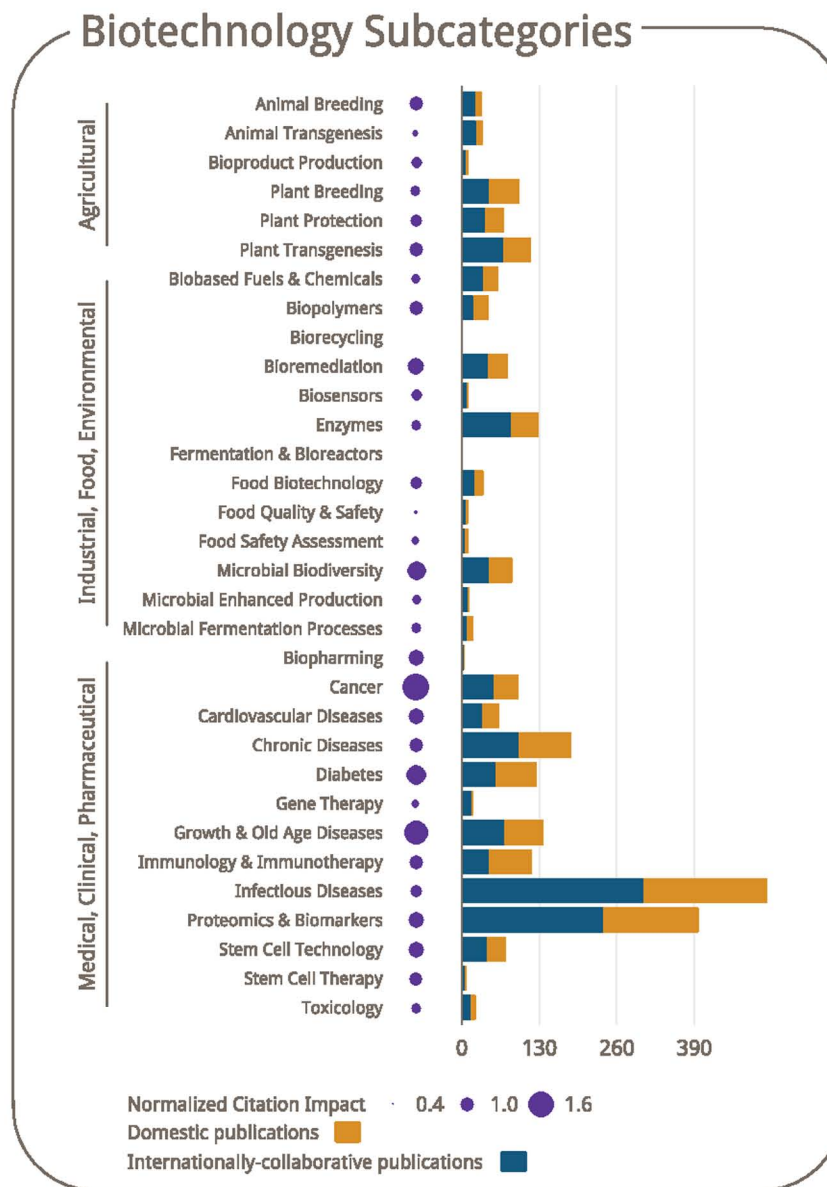
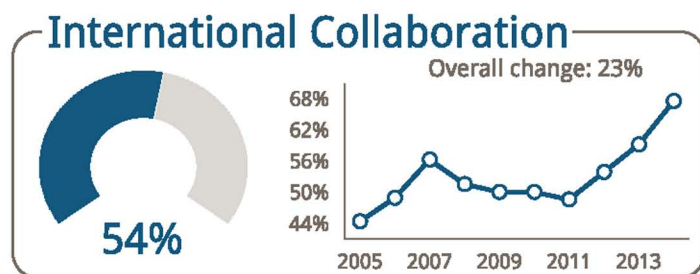
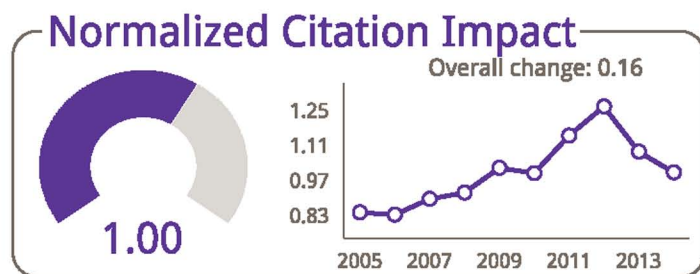
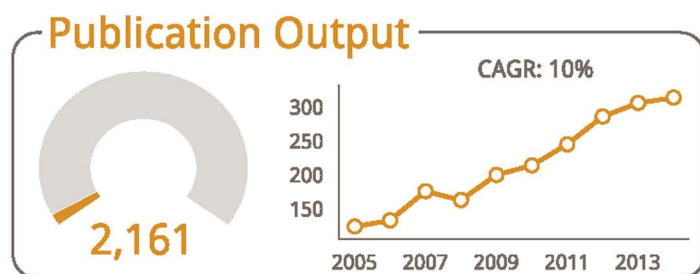


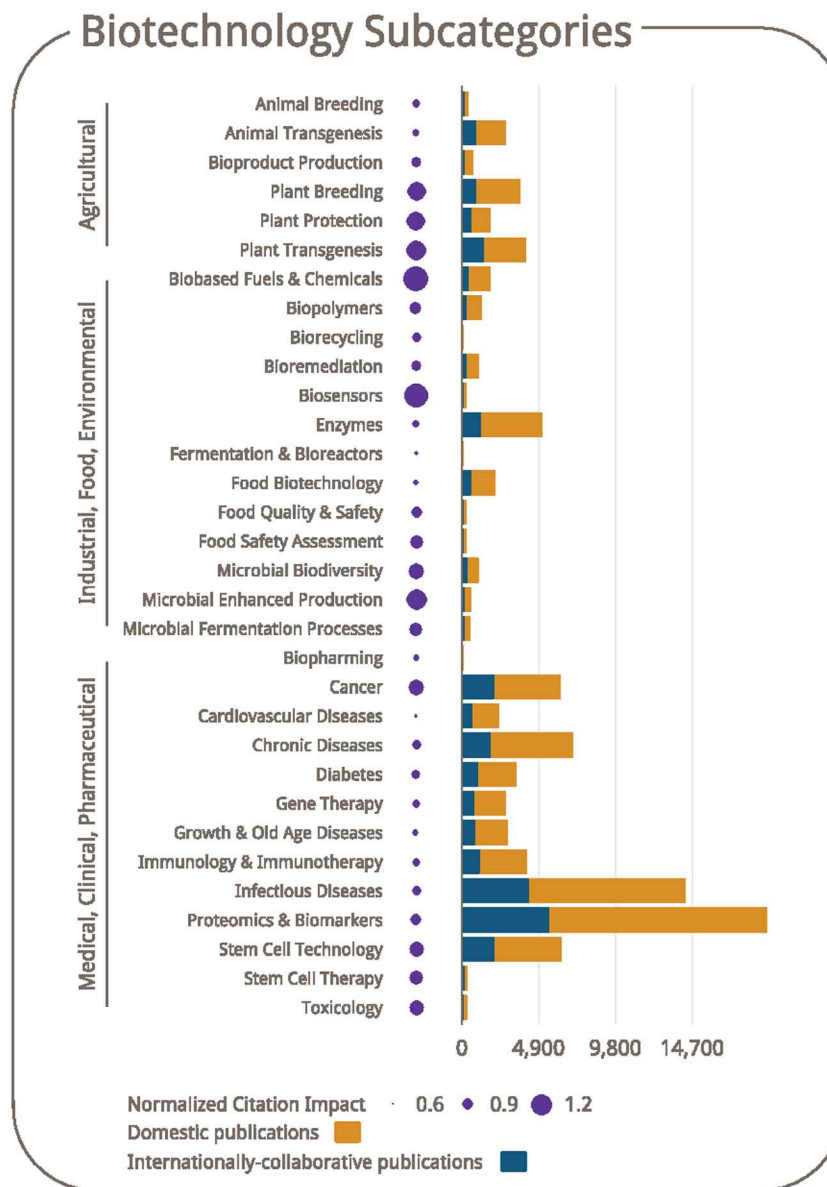
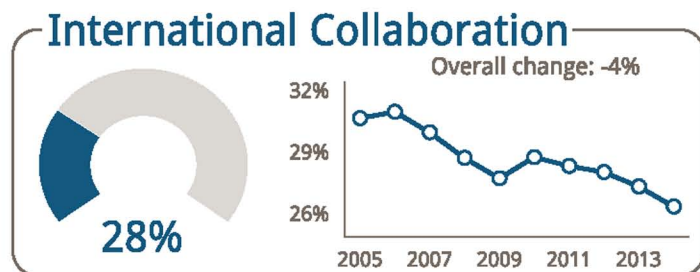
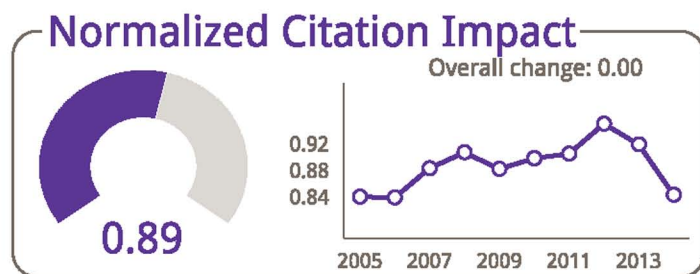
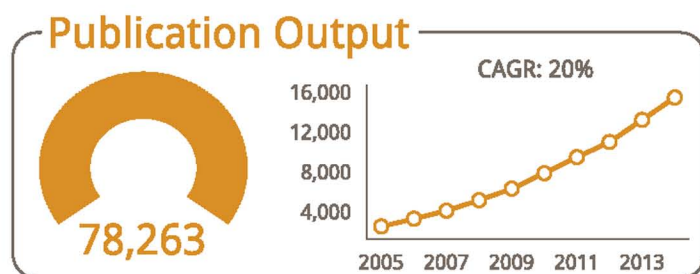


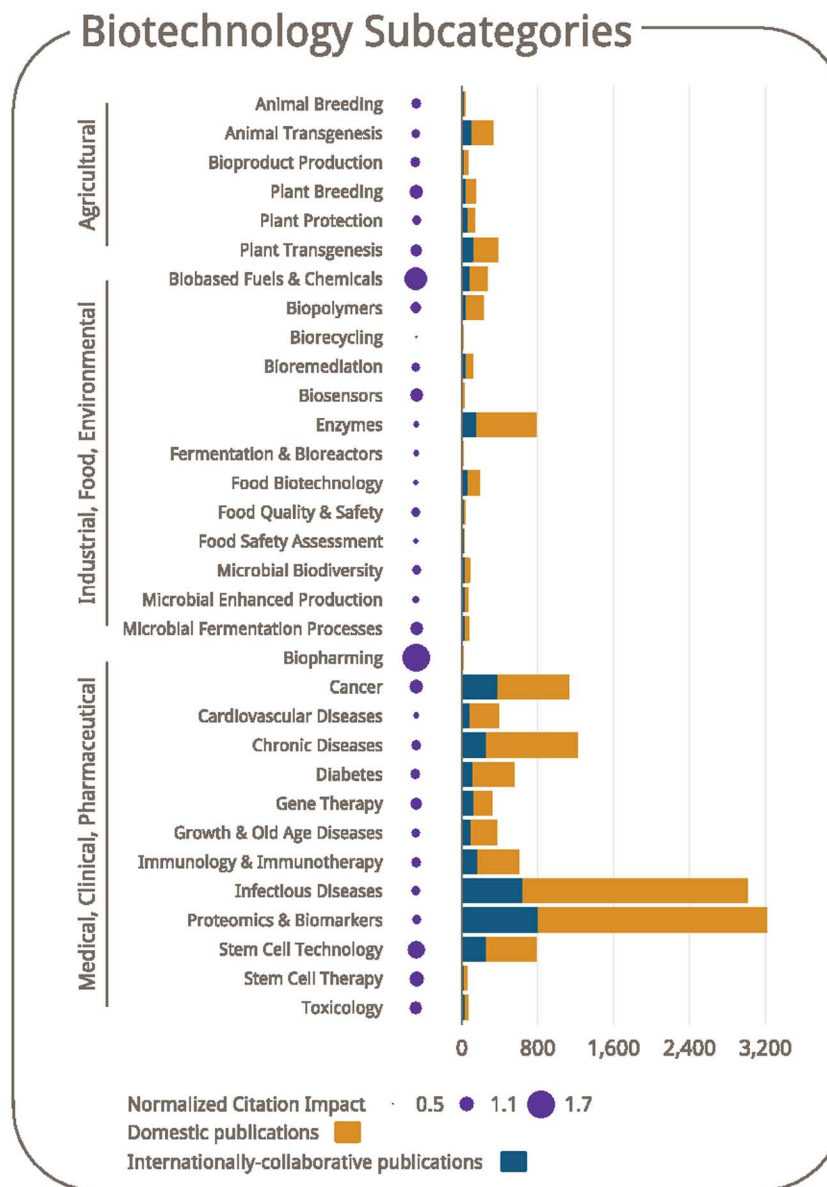
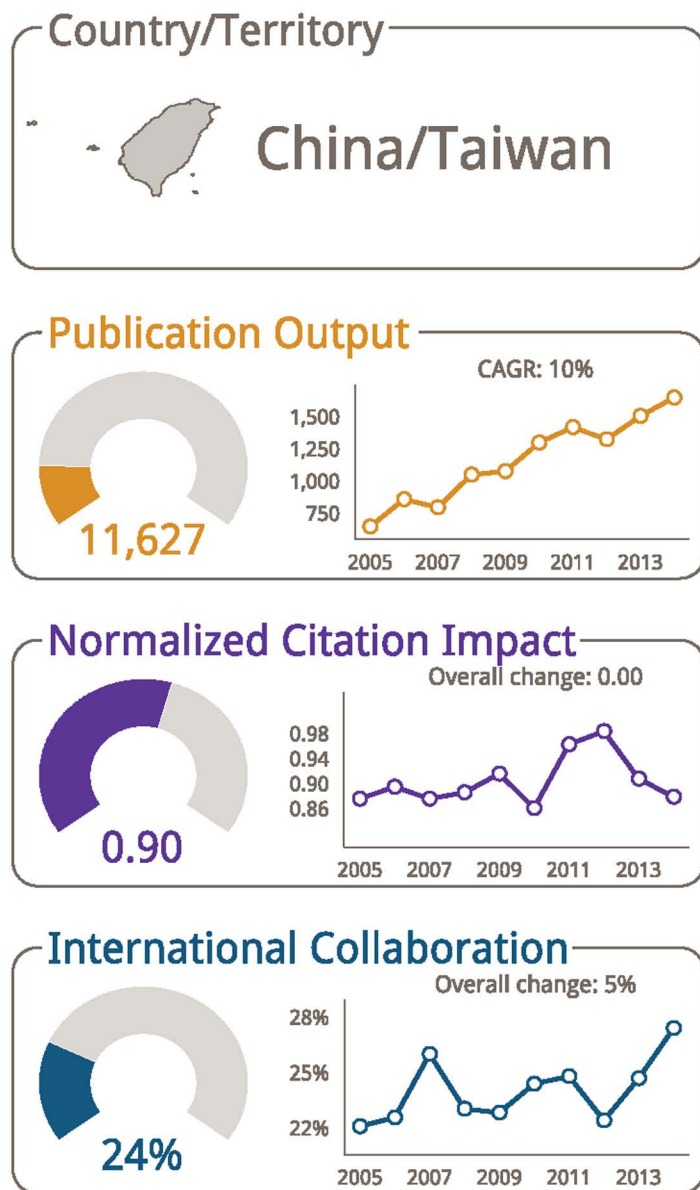


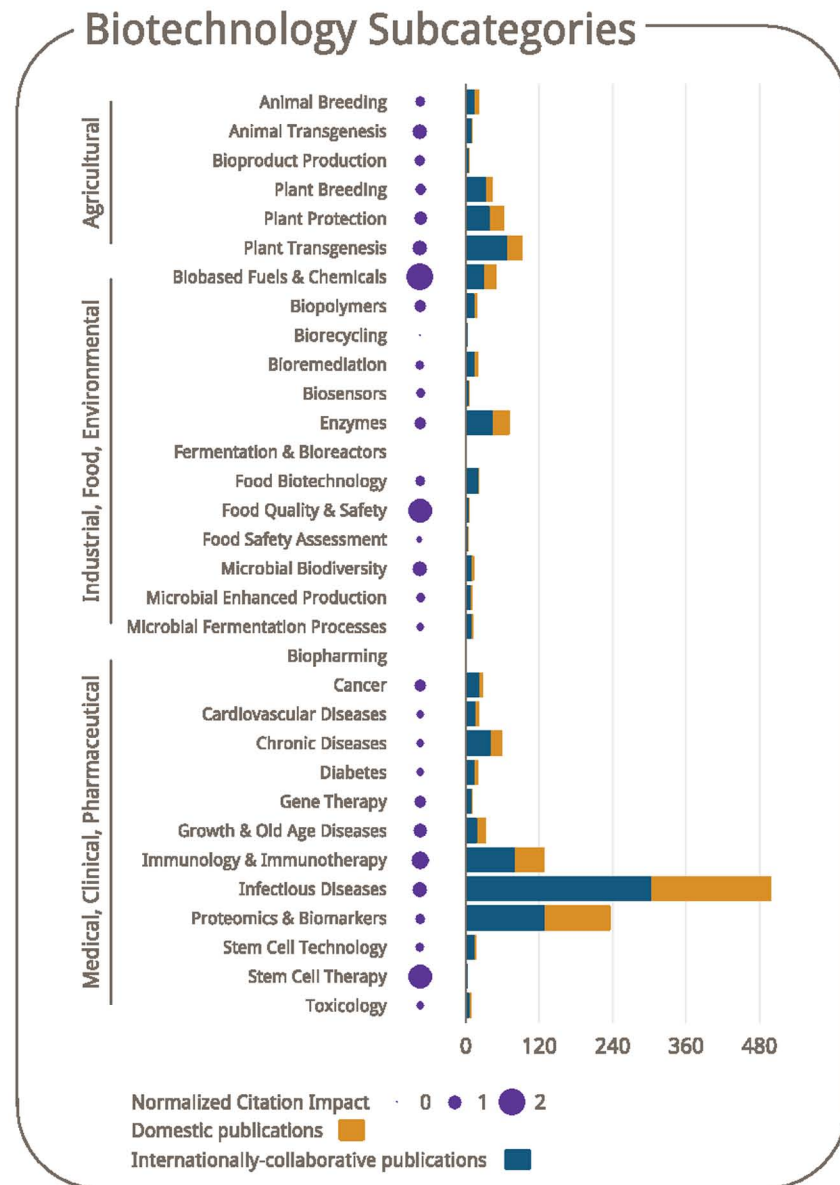
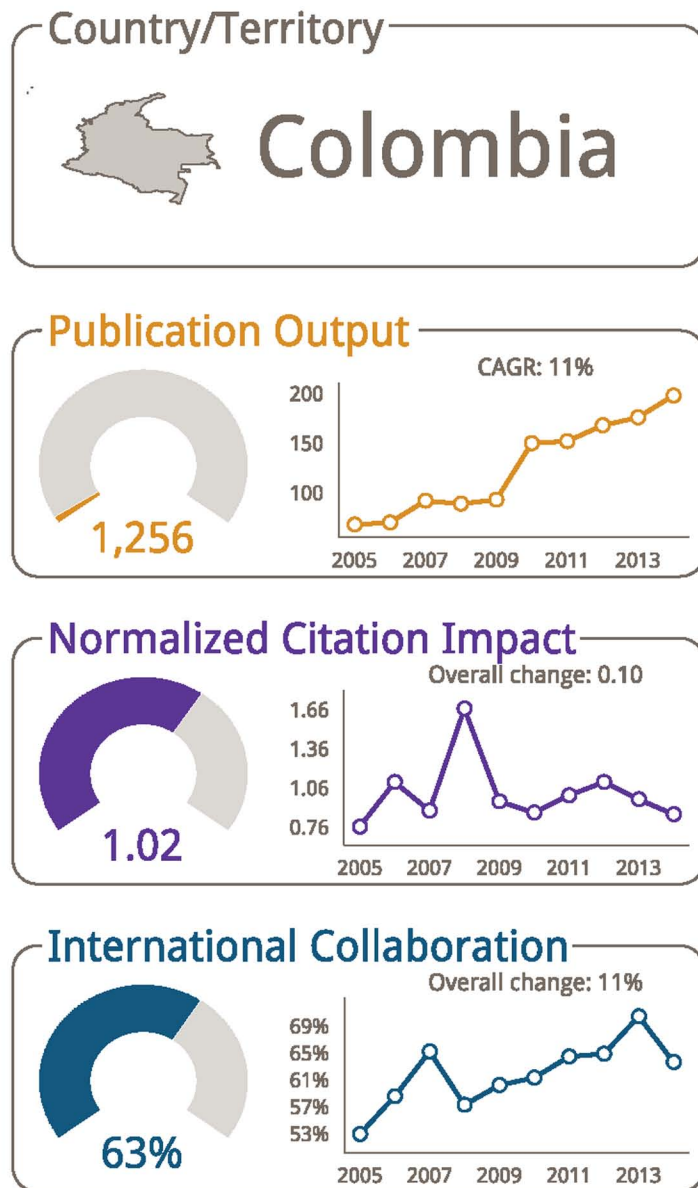




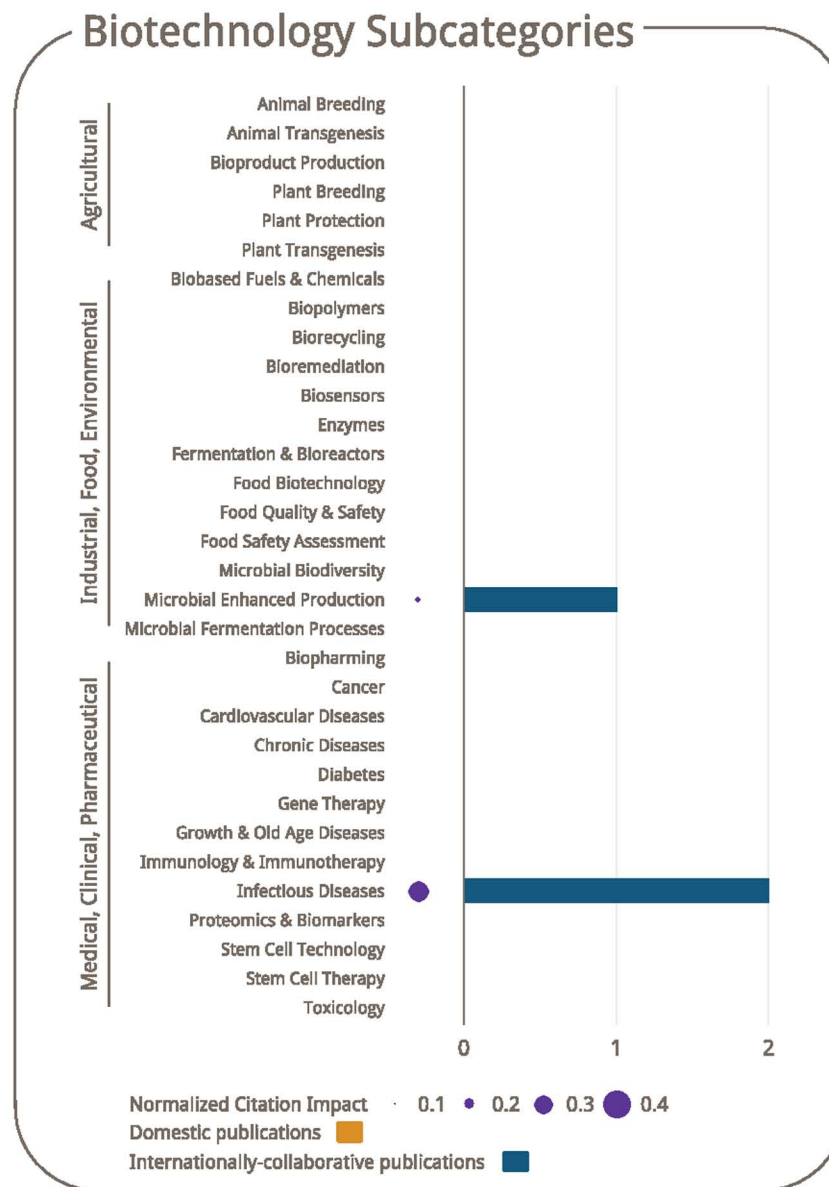
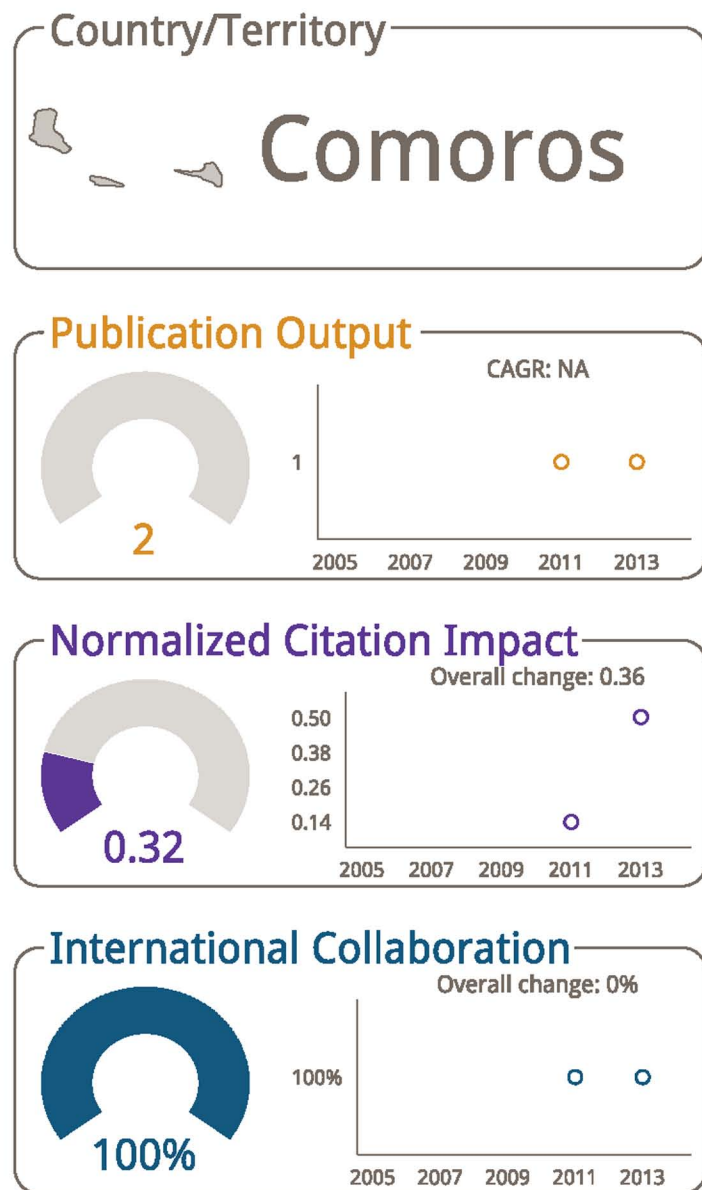


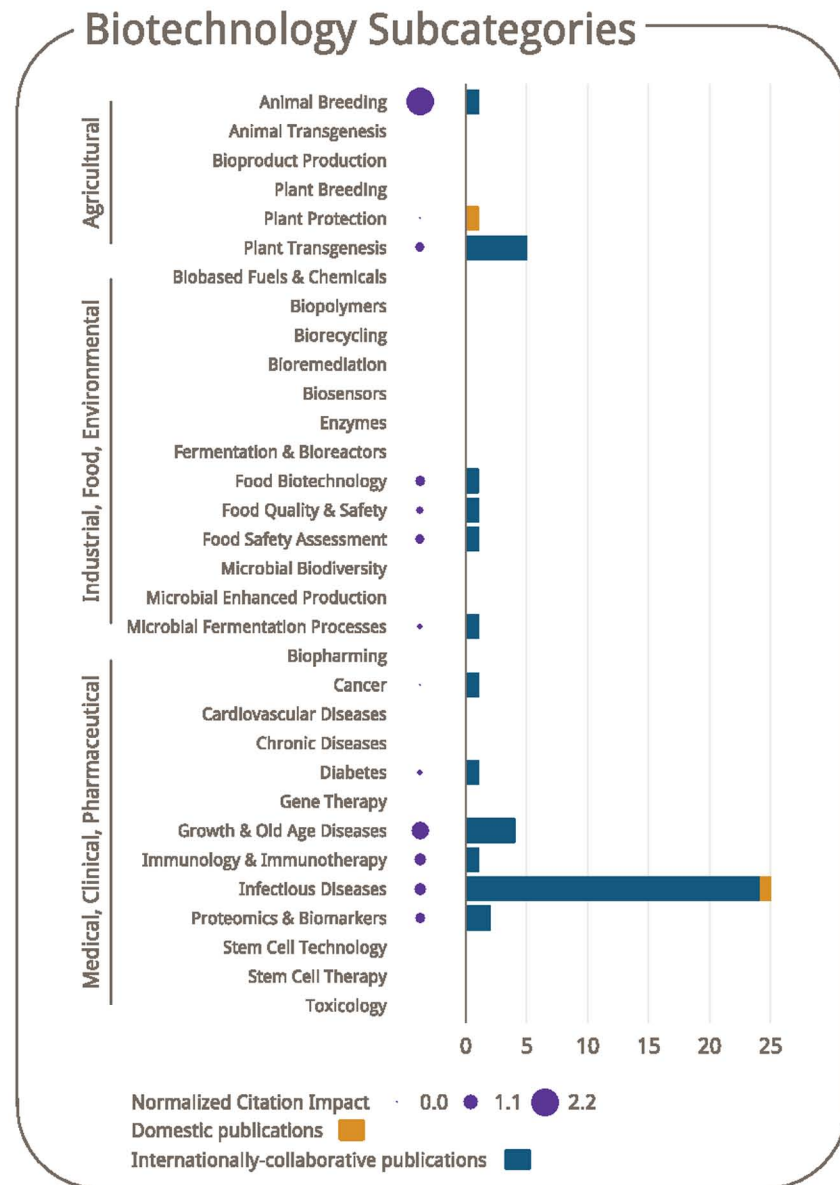
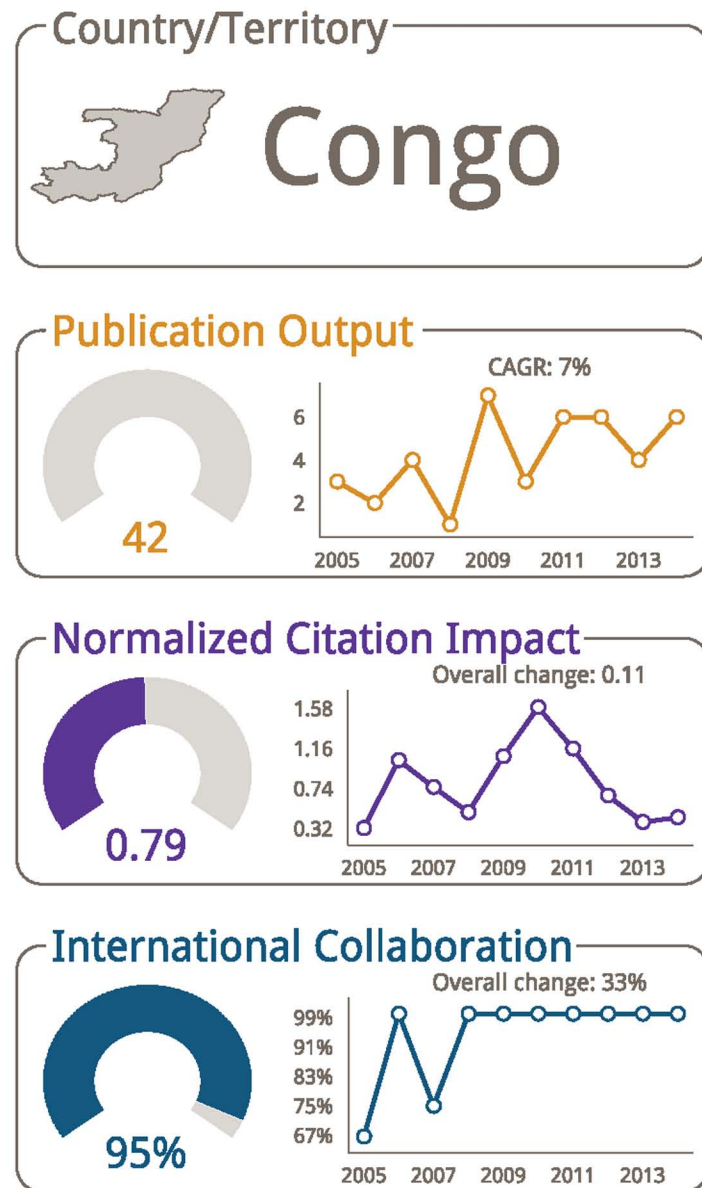


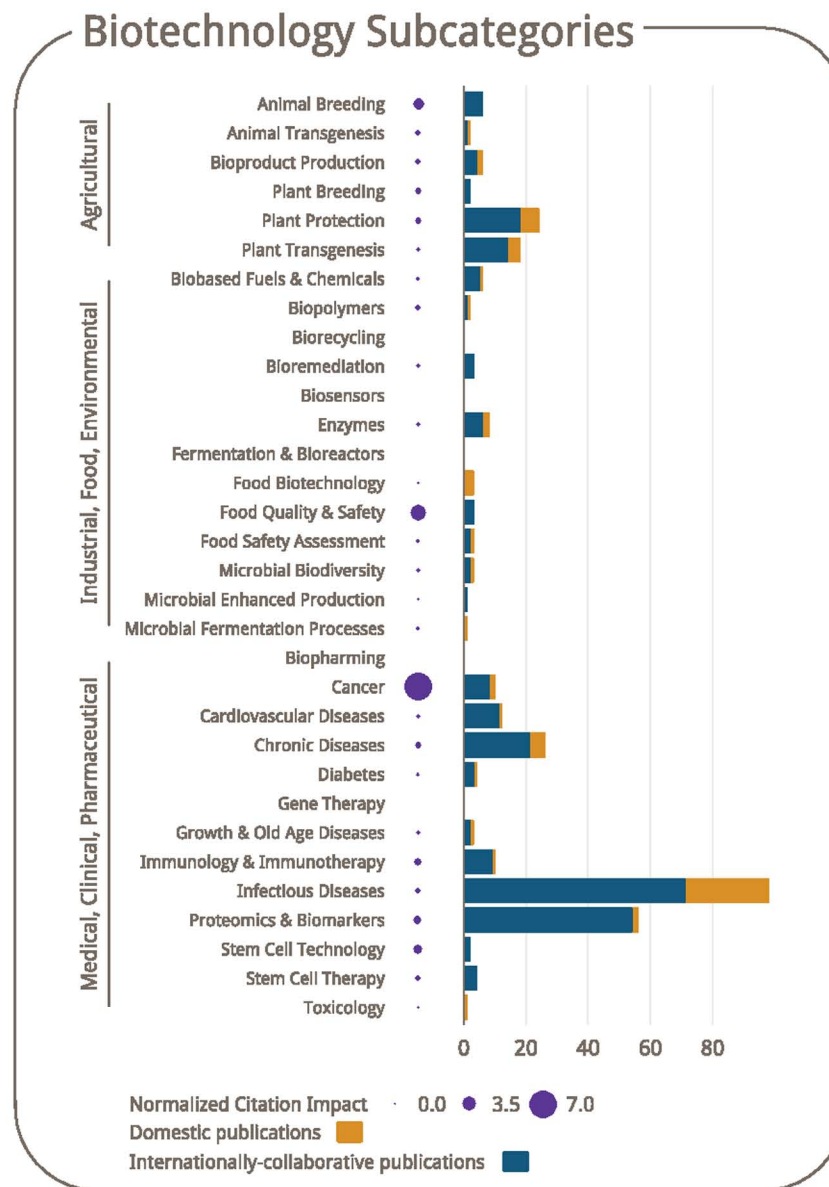
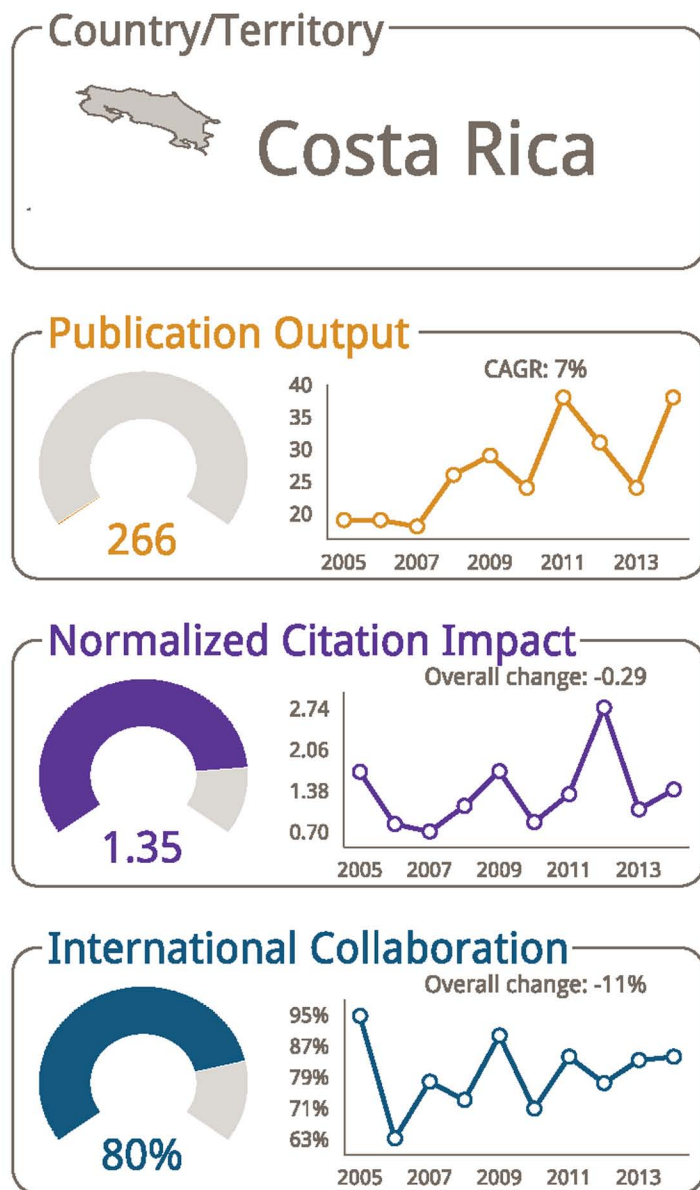




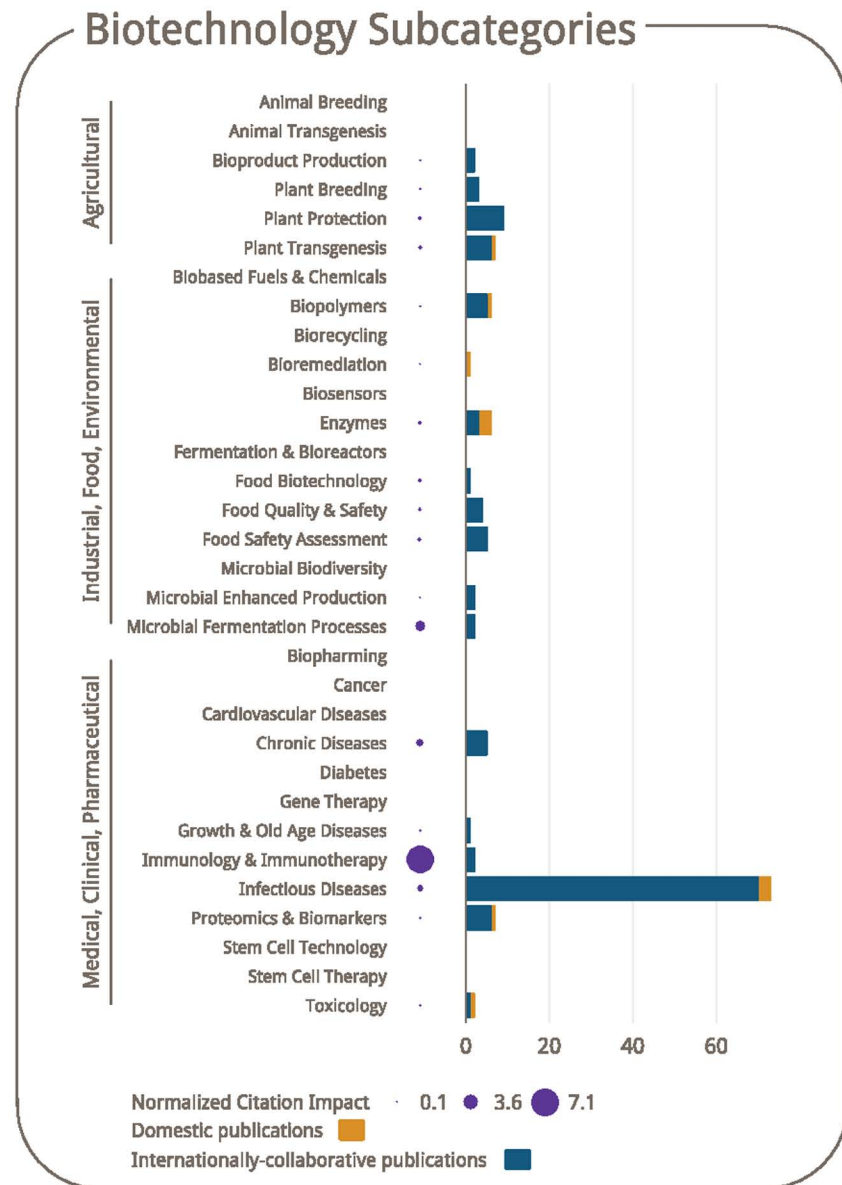
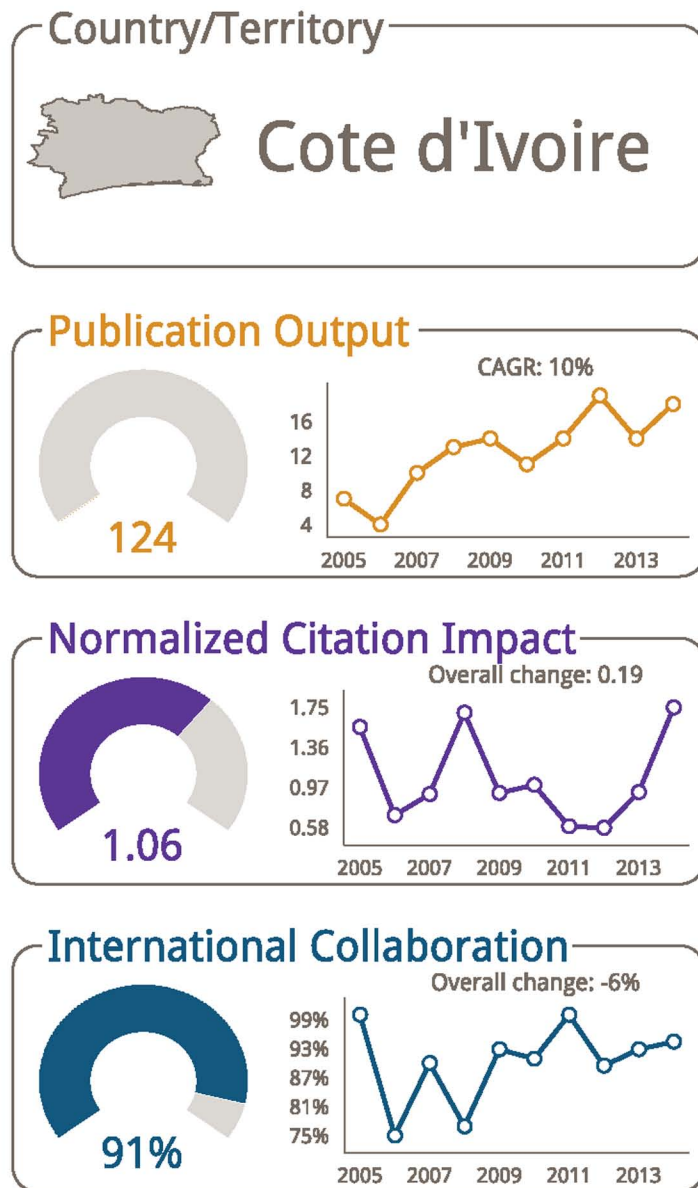


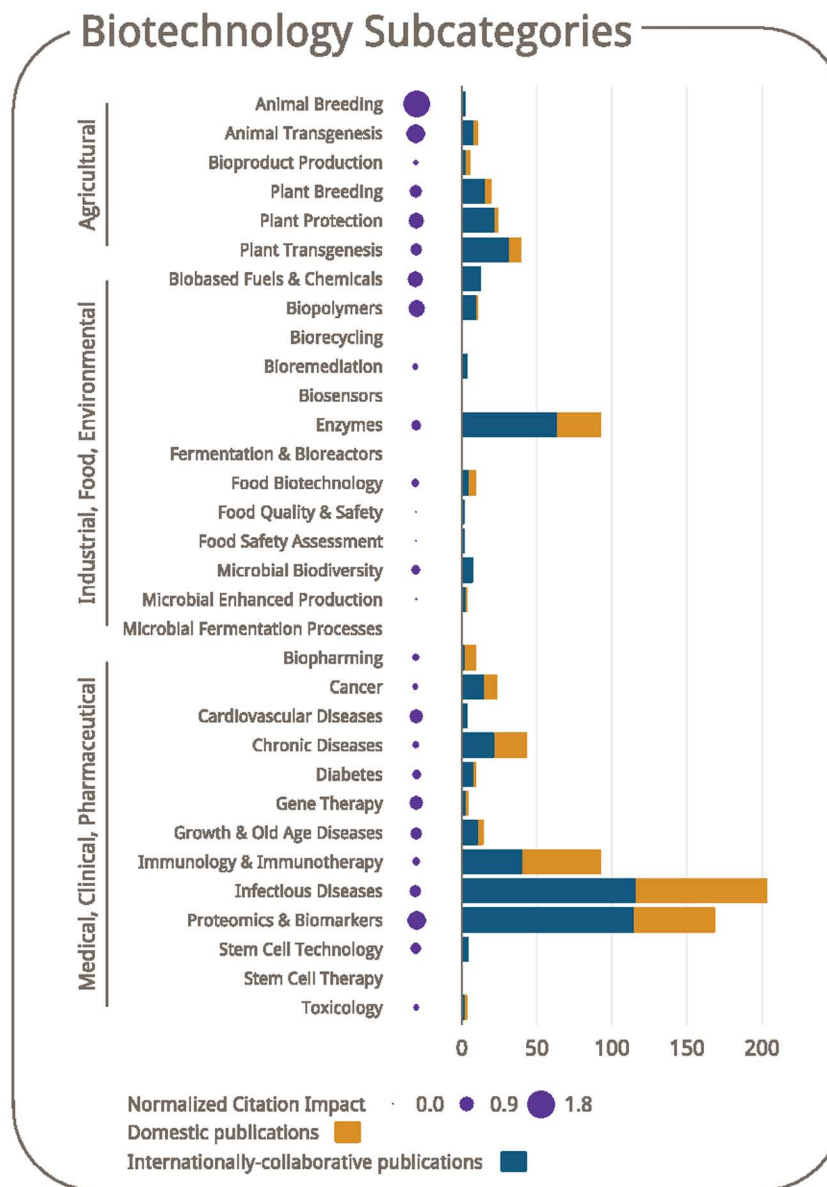
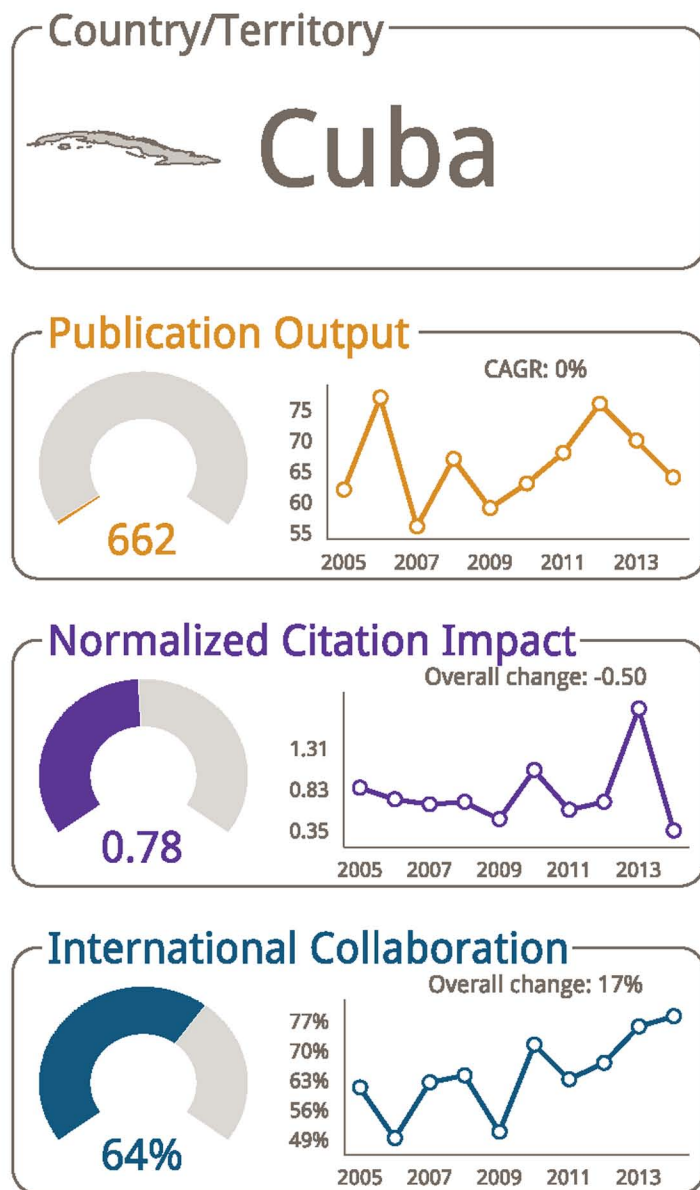












## Country/Territory

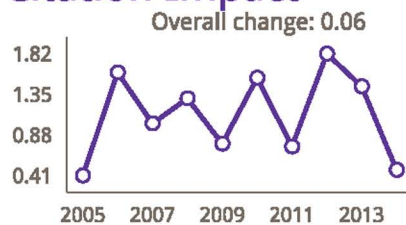


Democratic Republic of the Congo

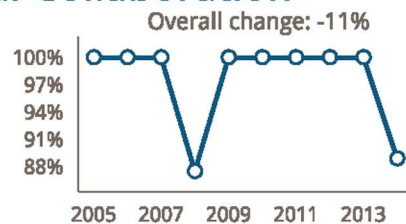
## Publication Output



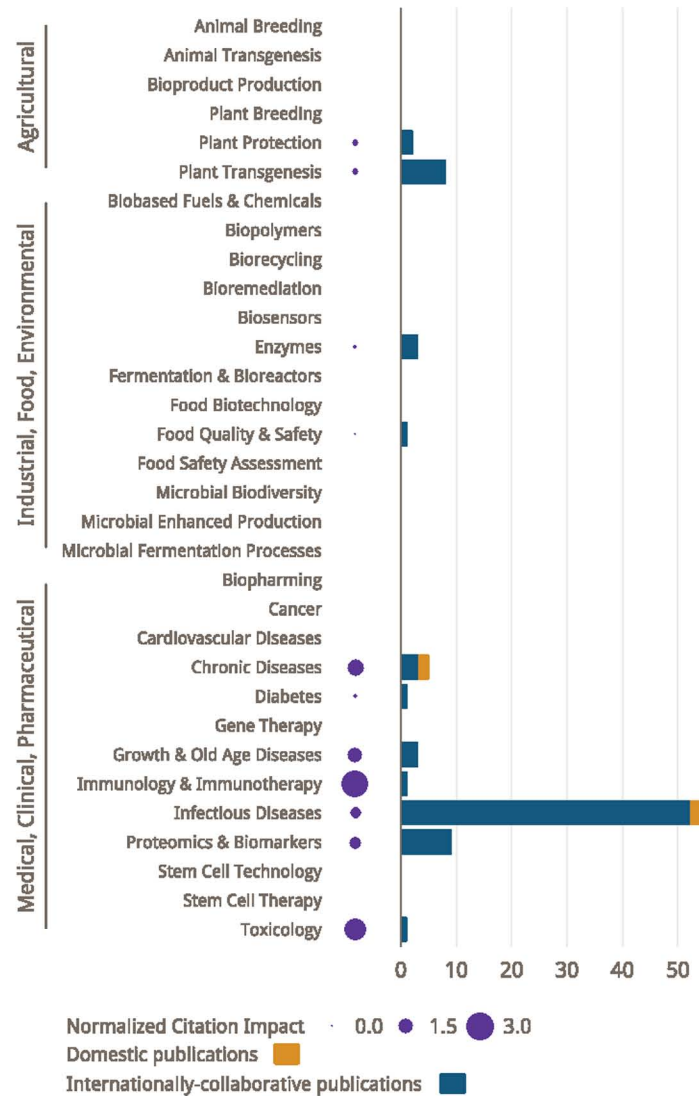
## Normalized Citation Impact

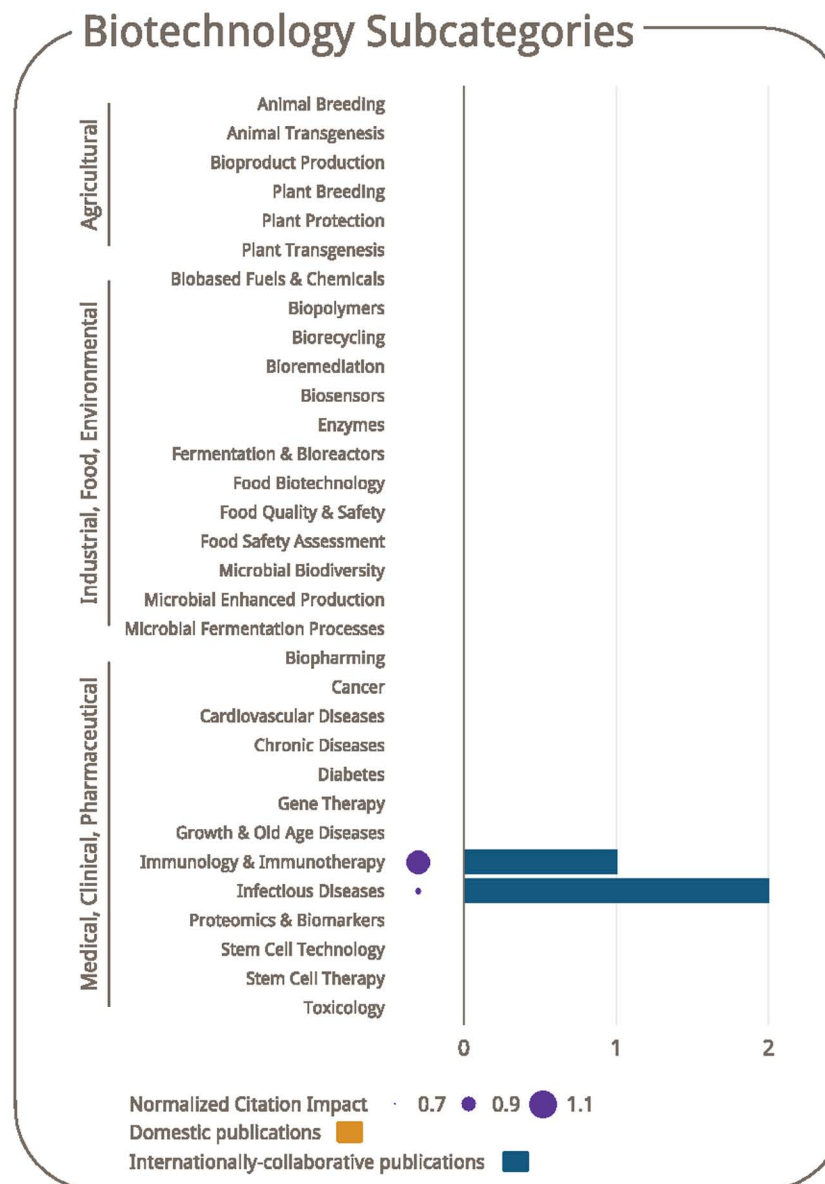
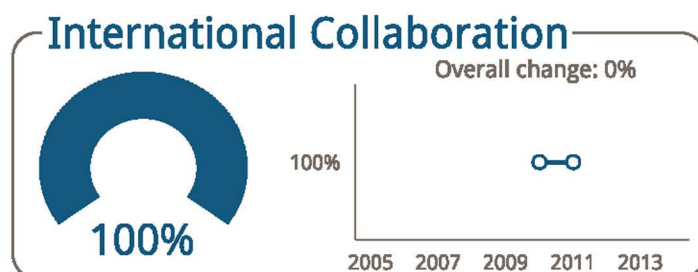
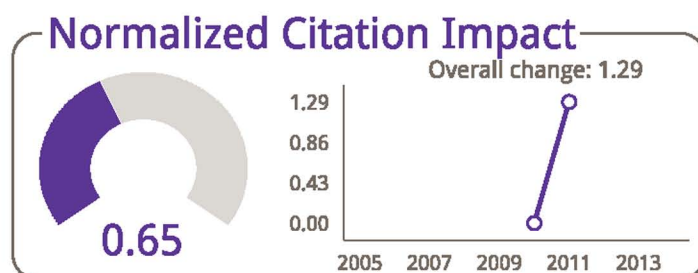
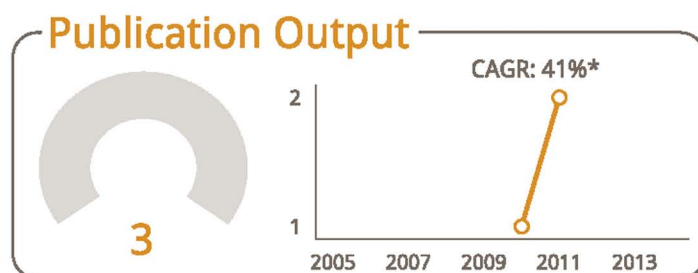


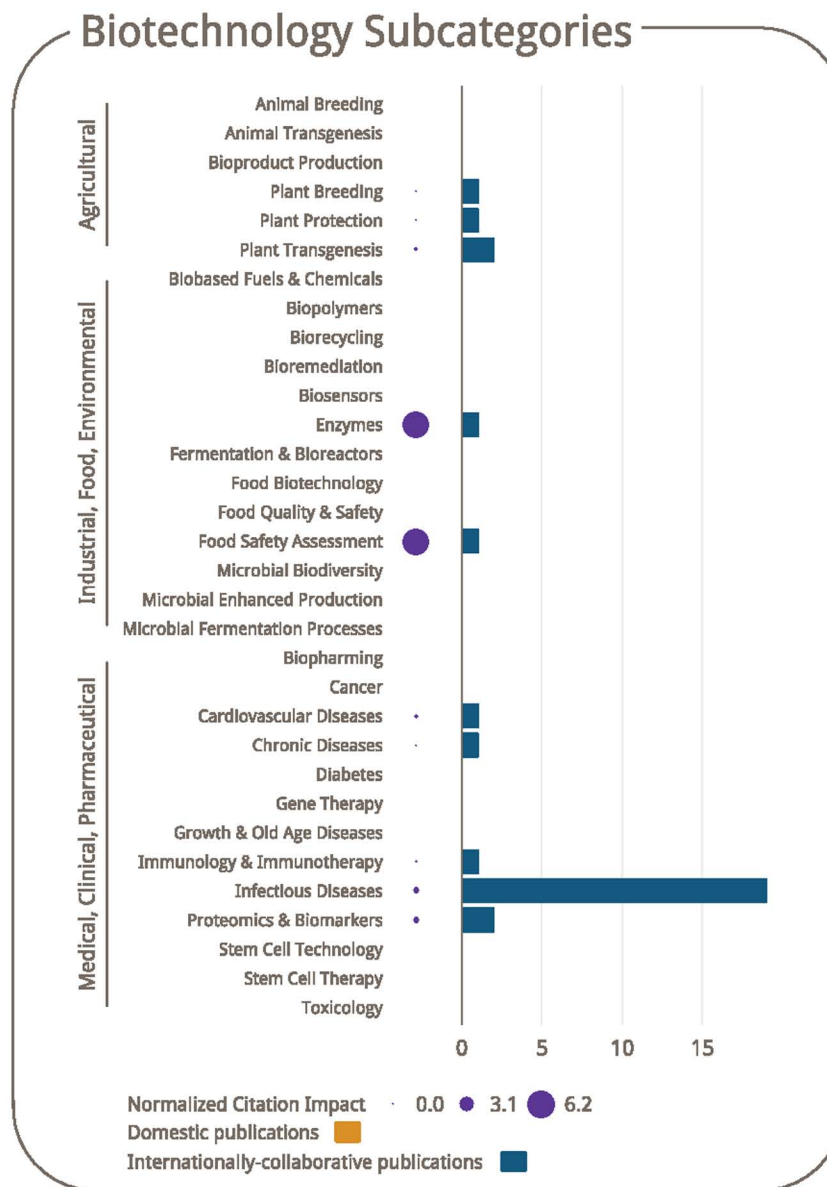
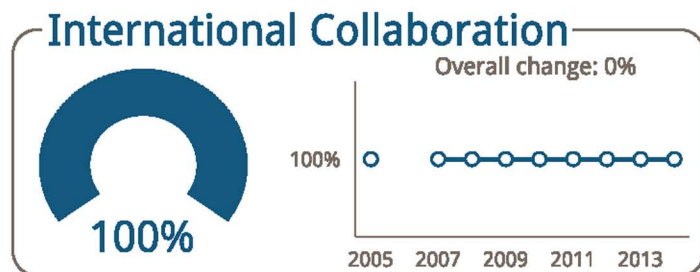
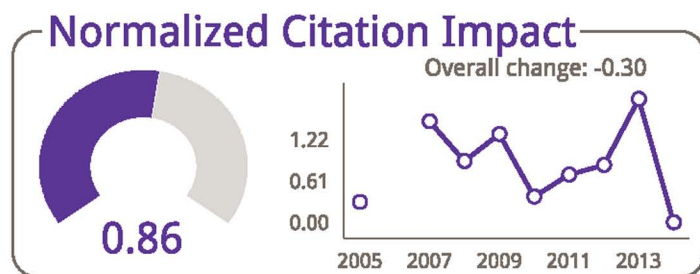
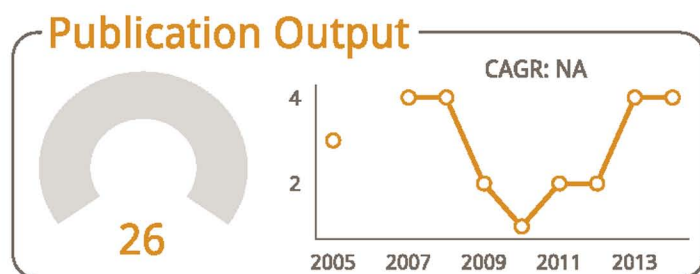
## International Collaboration



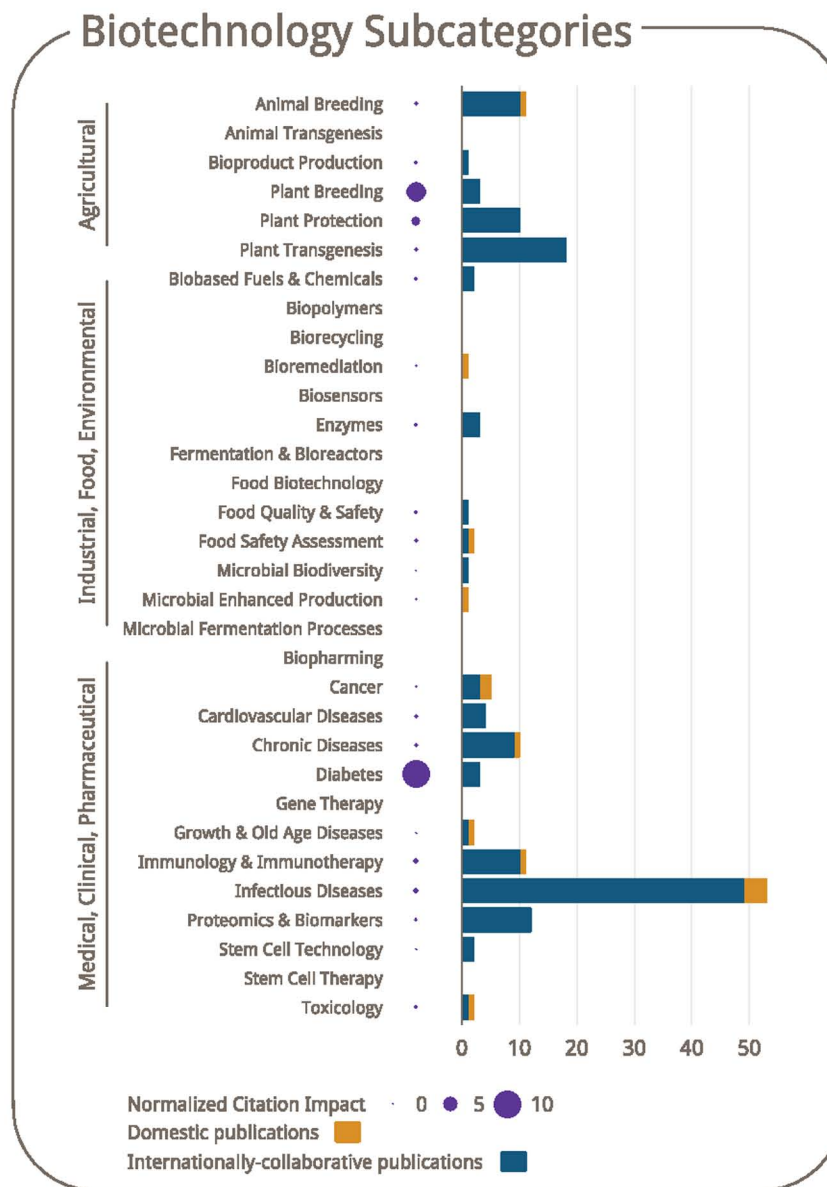
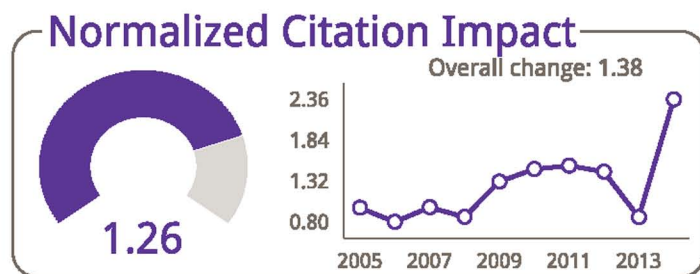
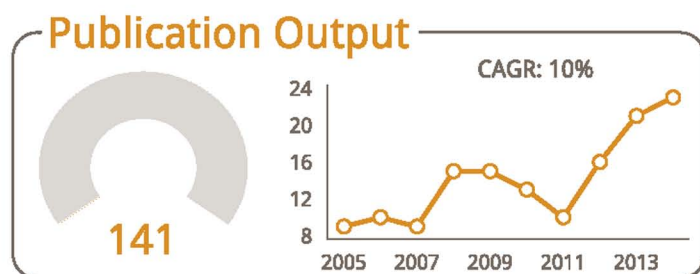
## Biotechnology Subcategories

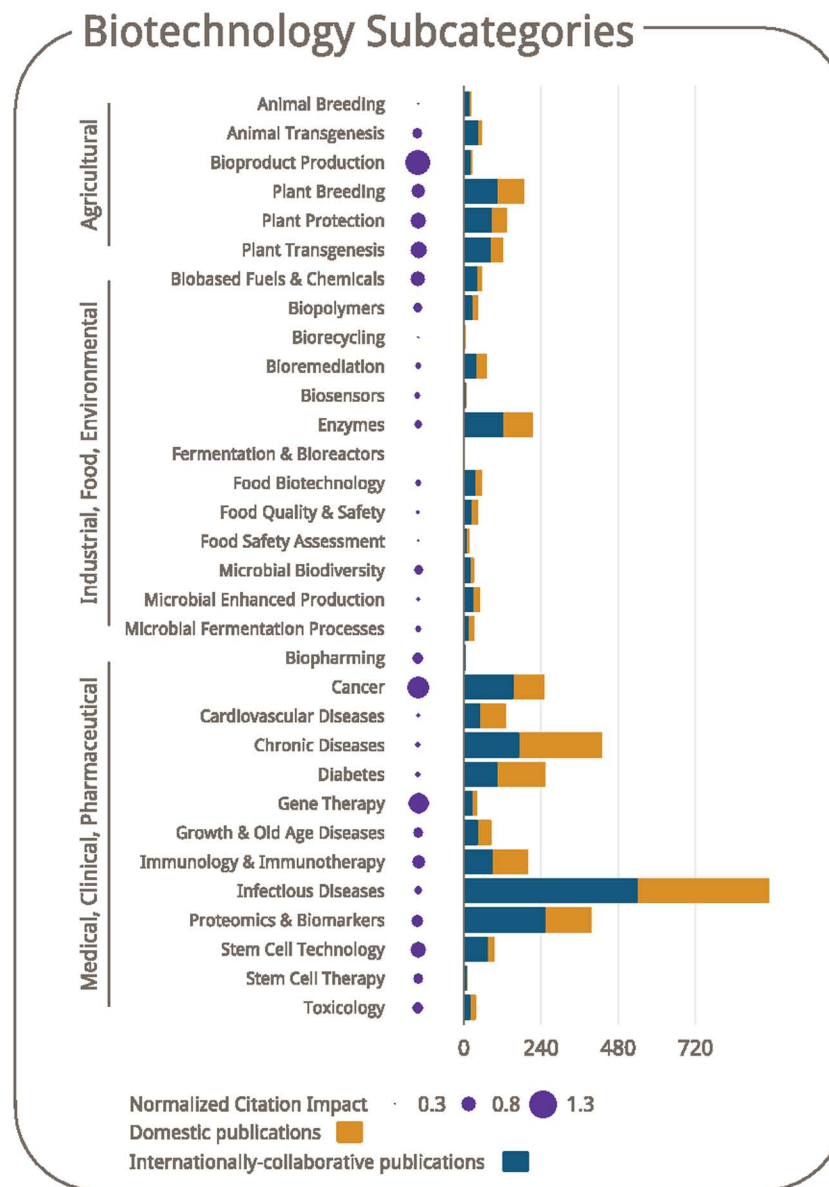
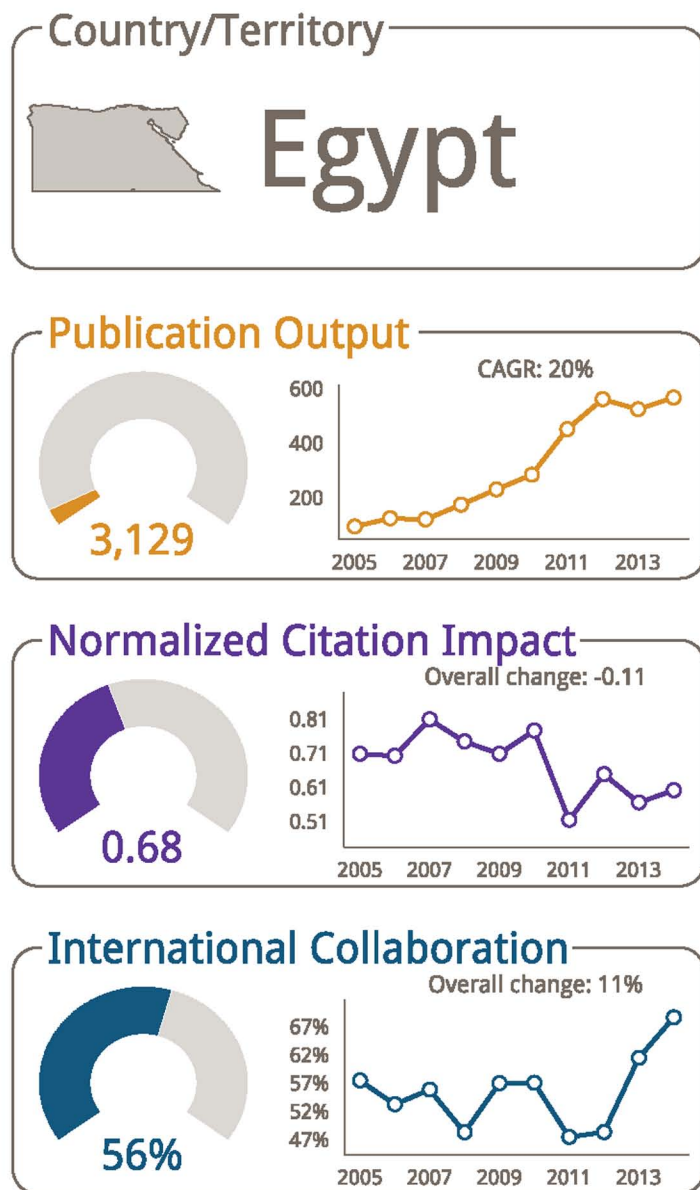


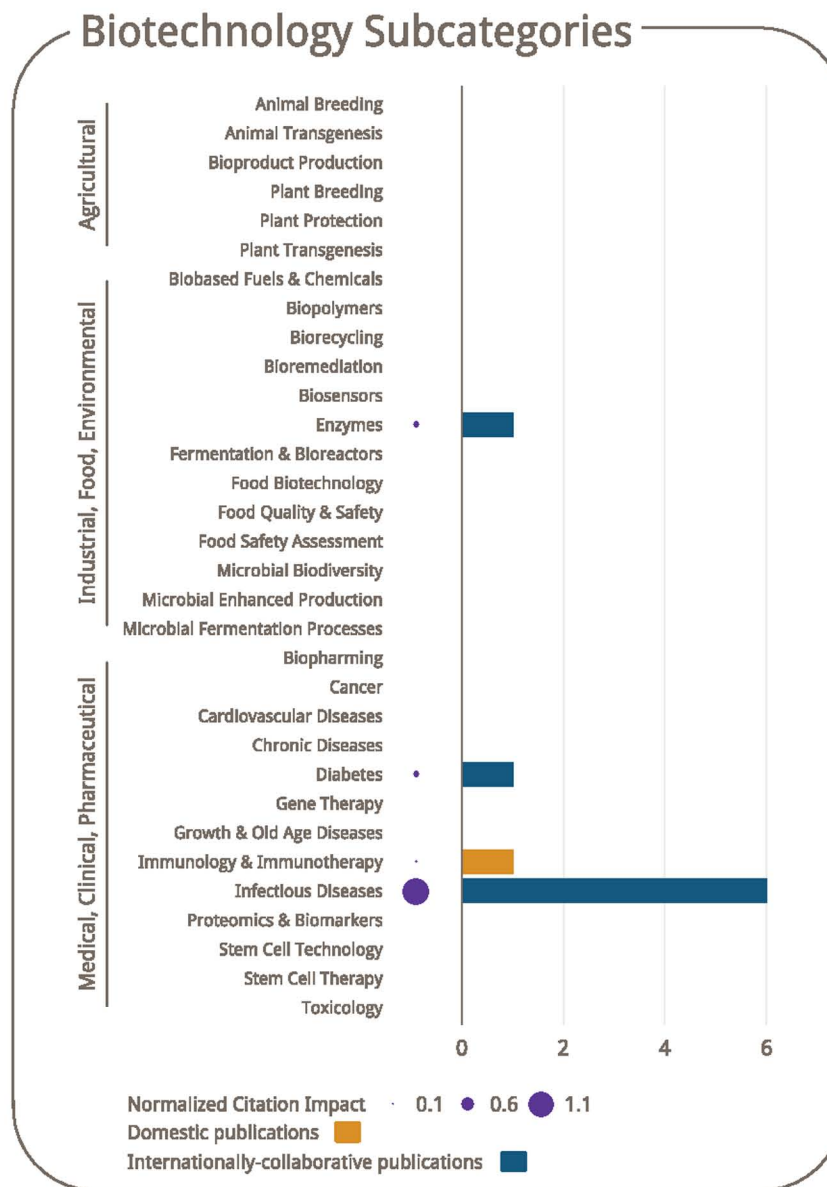
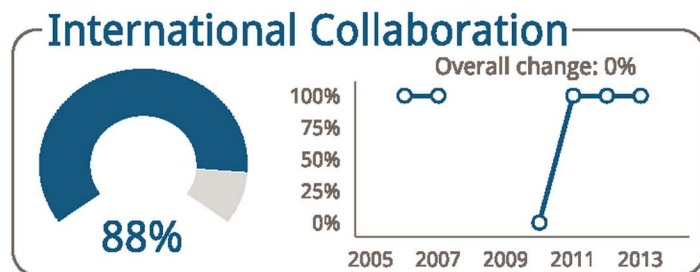
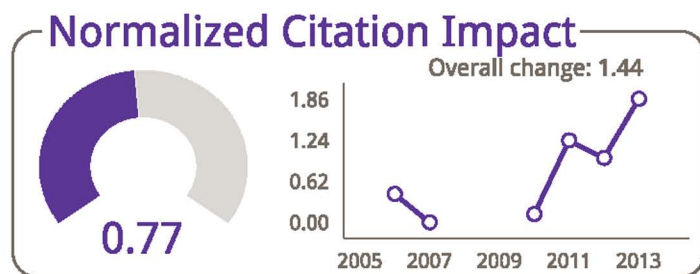
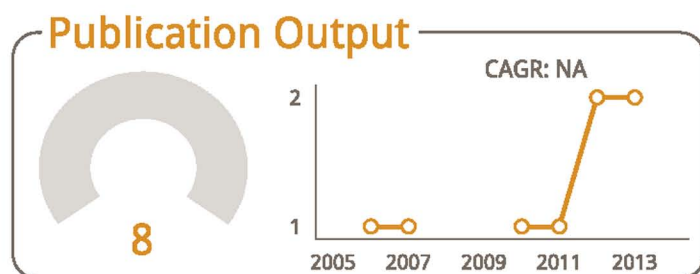




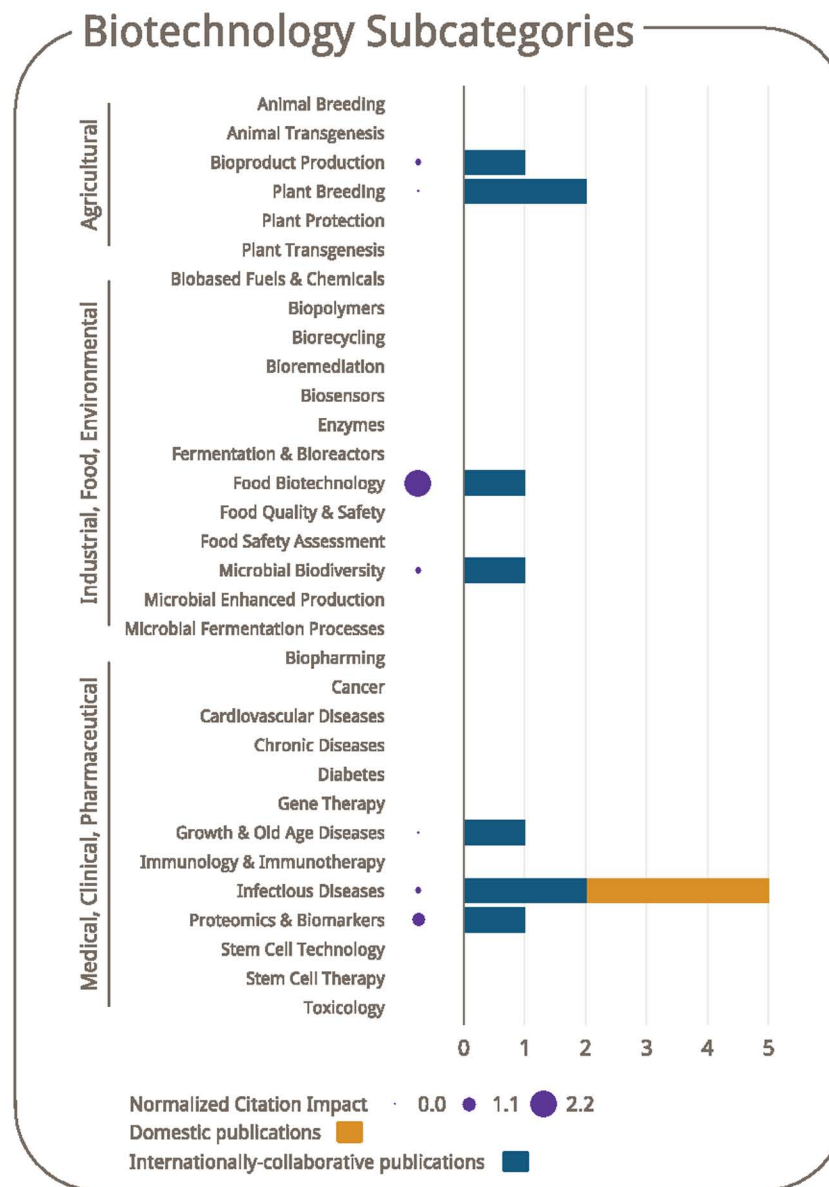
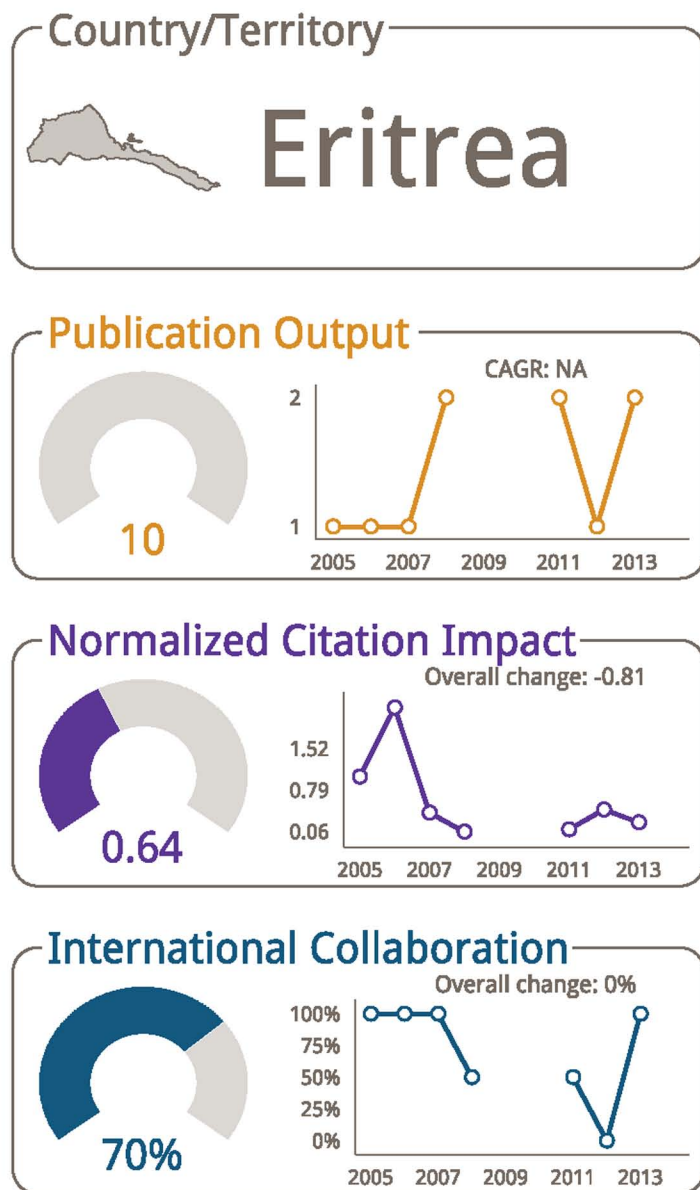


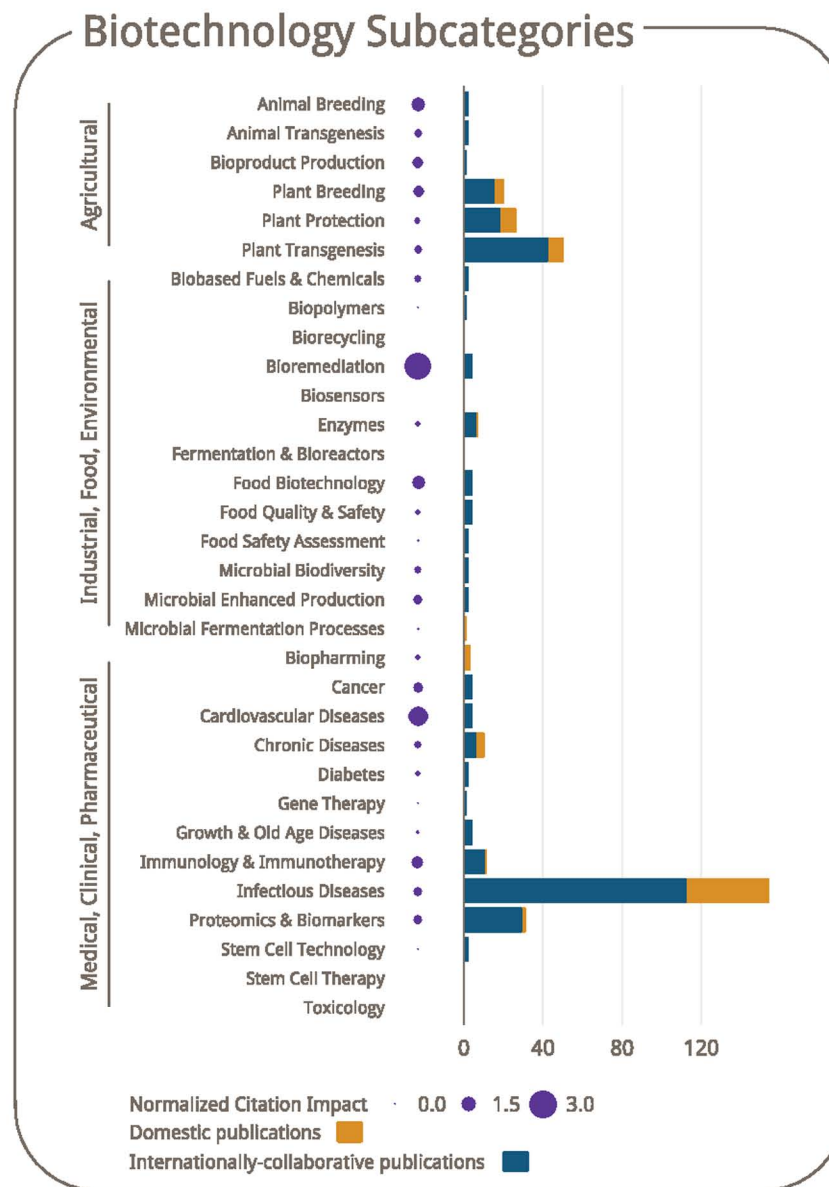
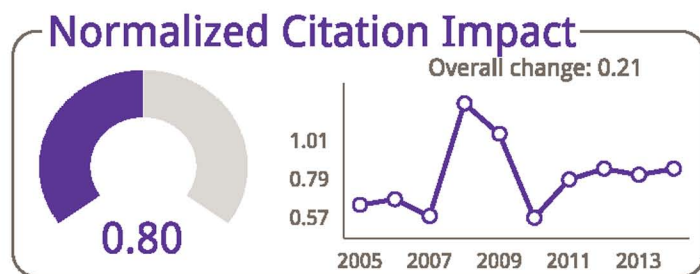
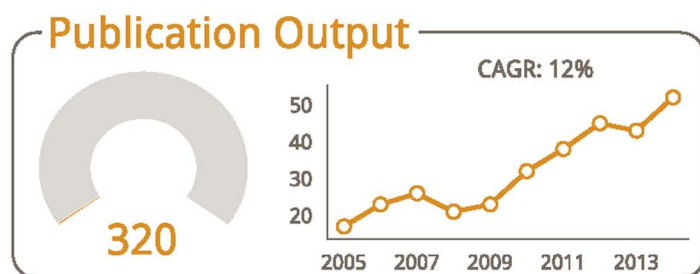


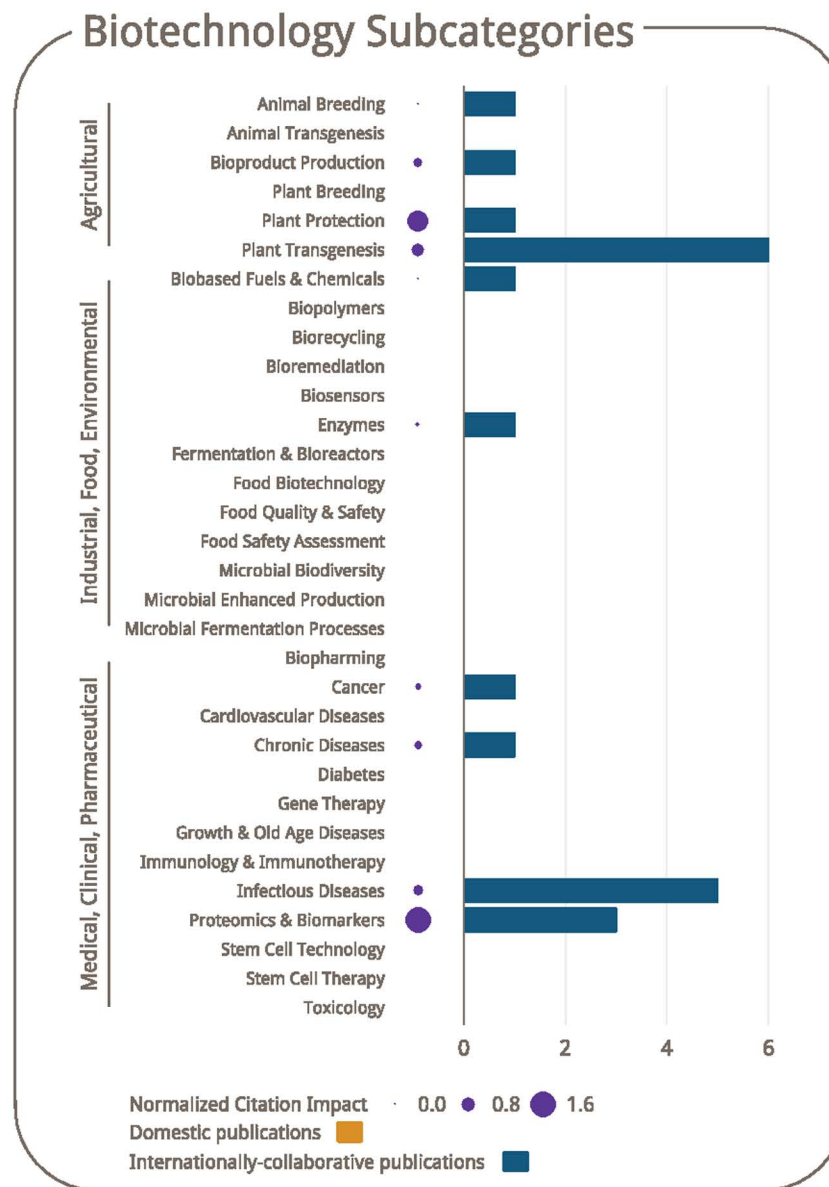
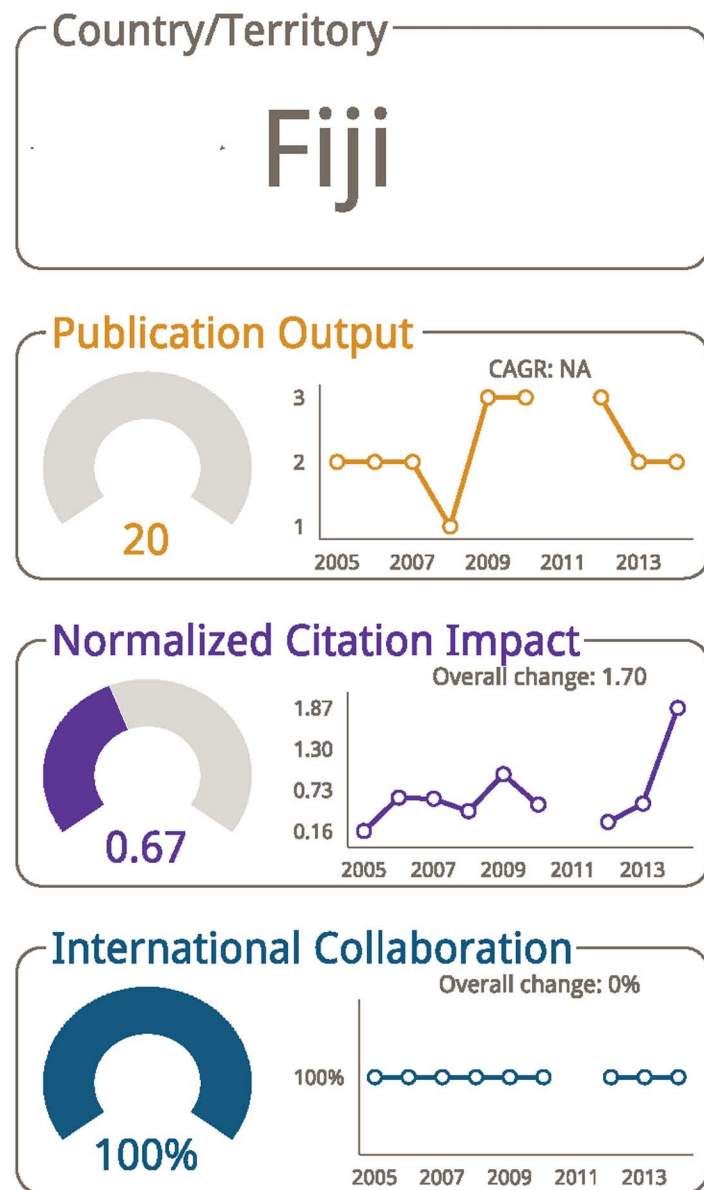


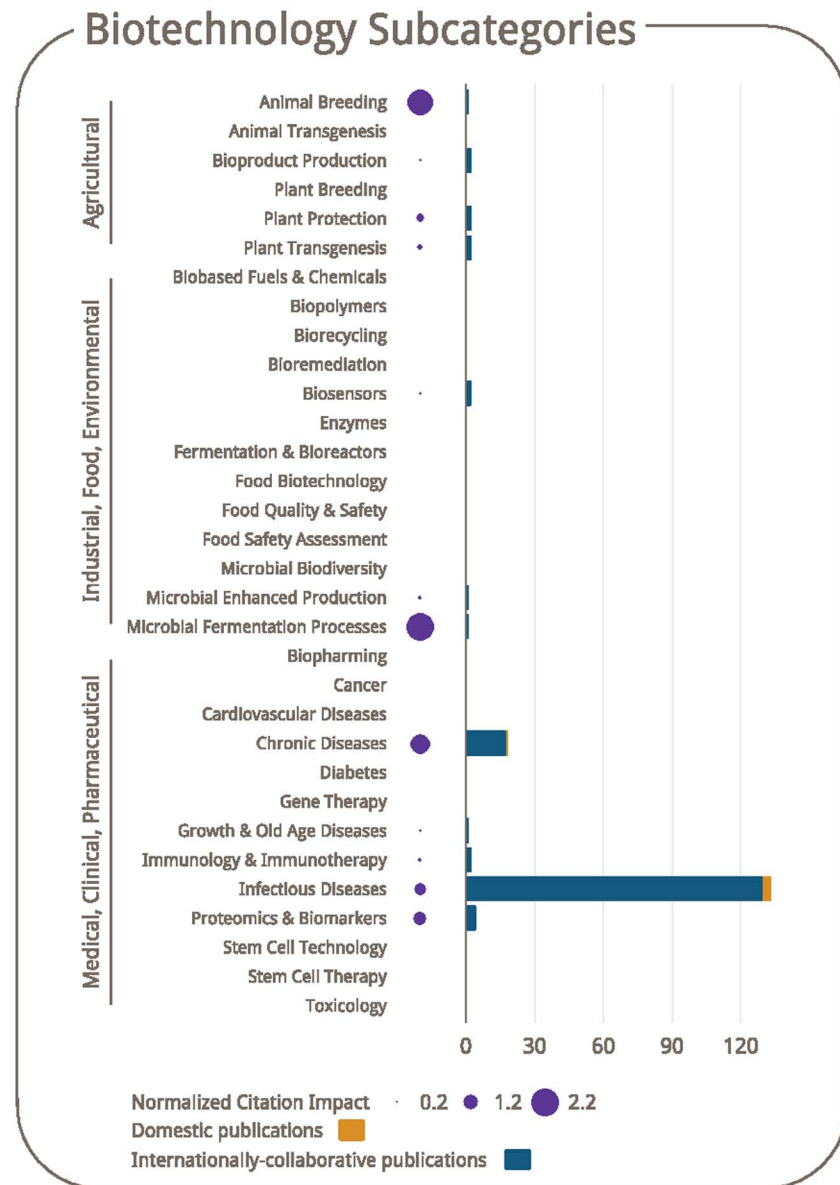
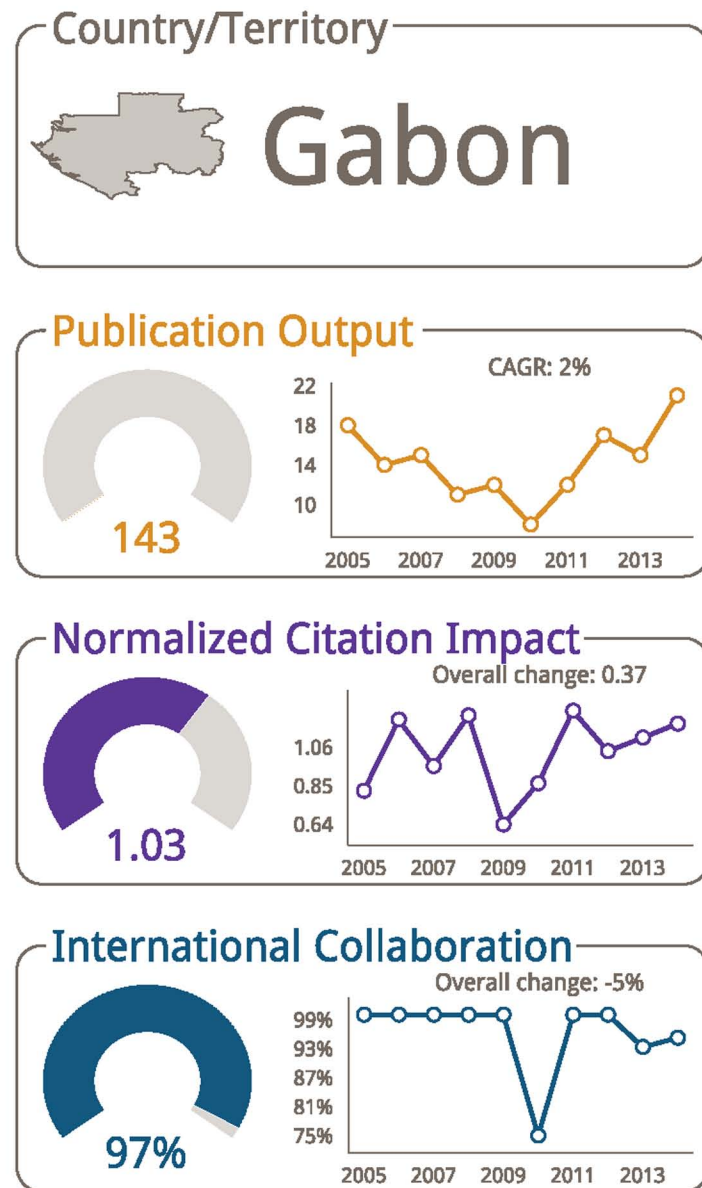


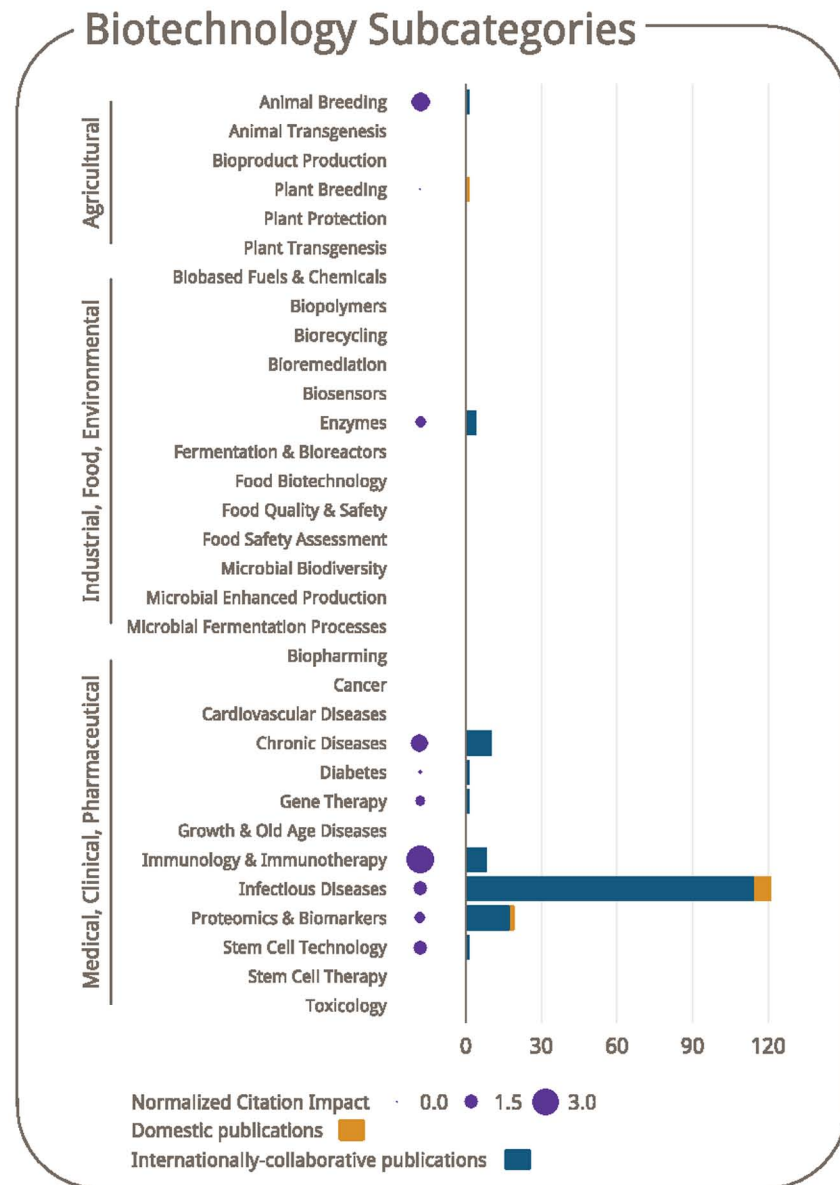
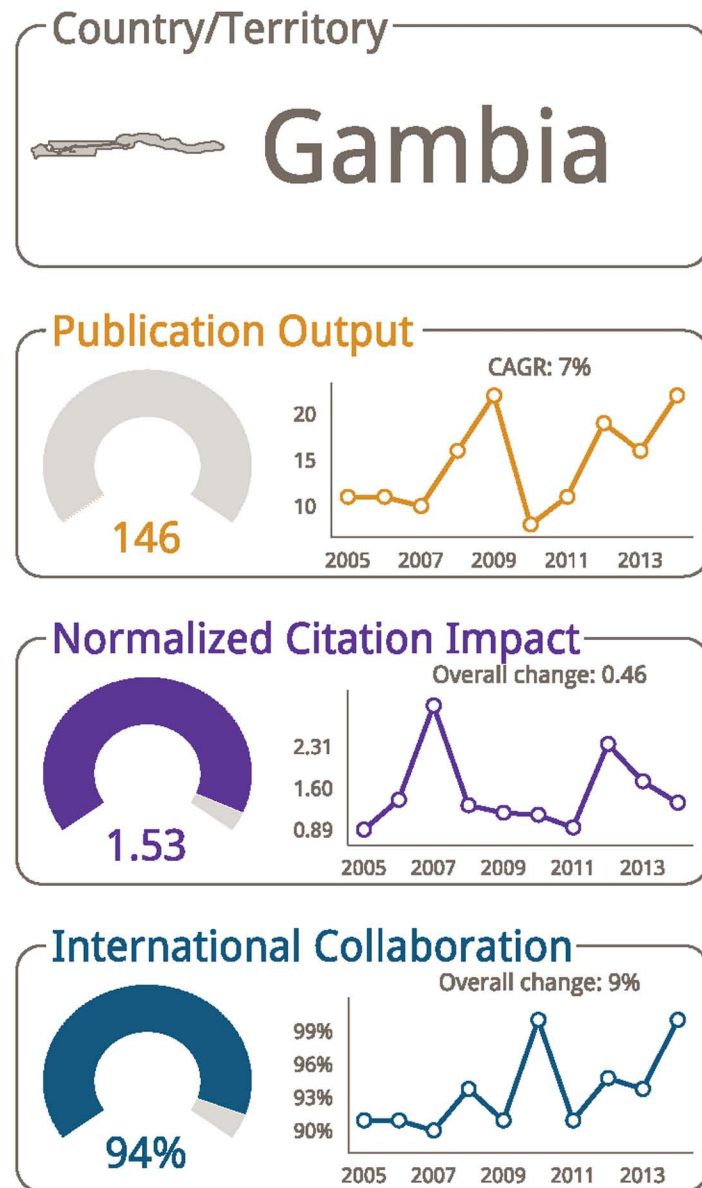




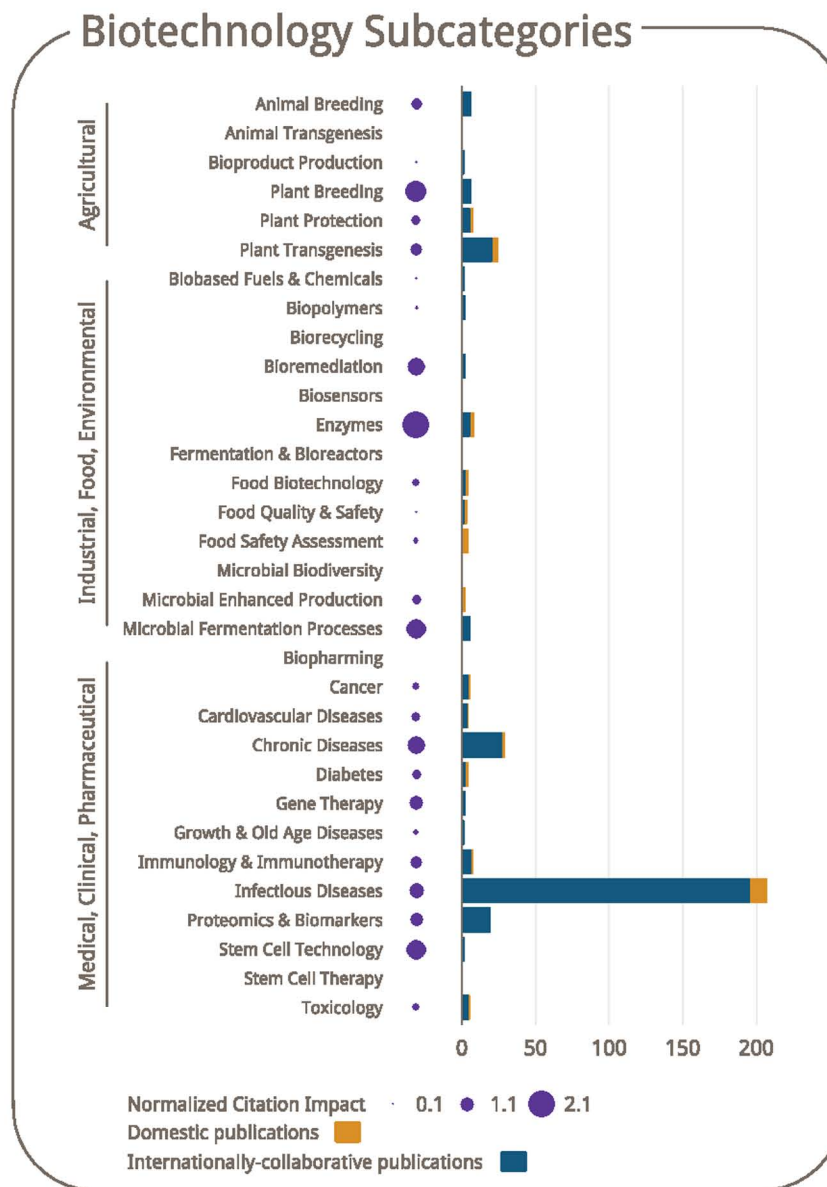
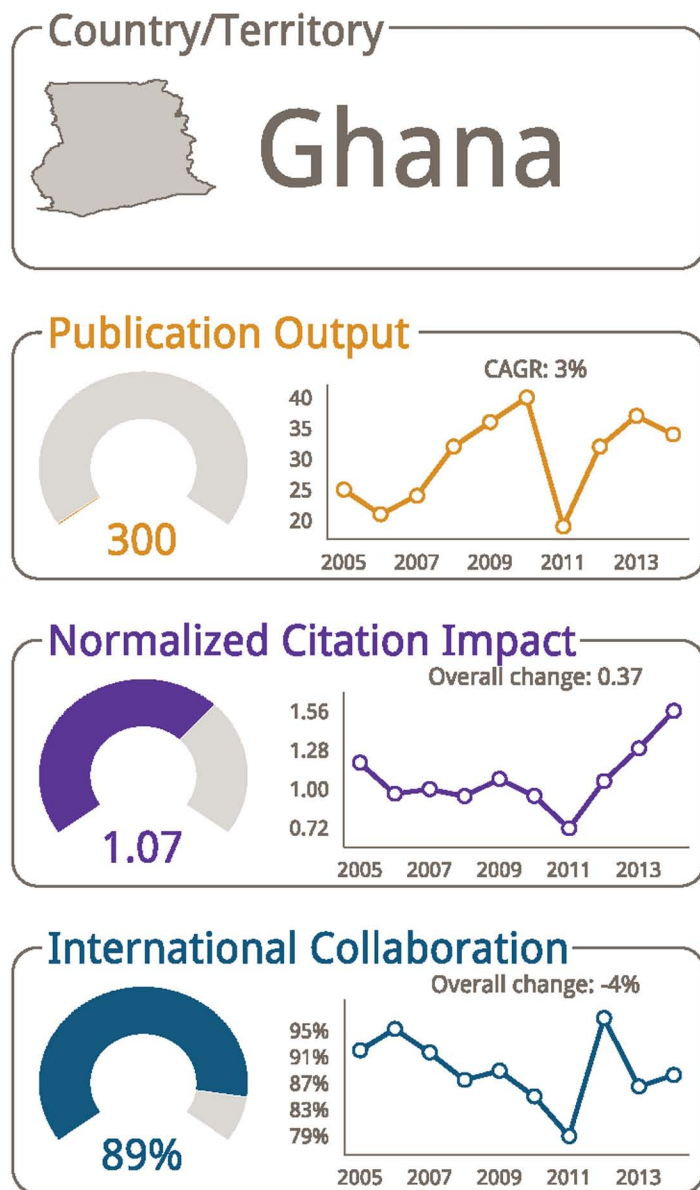


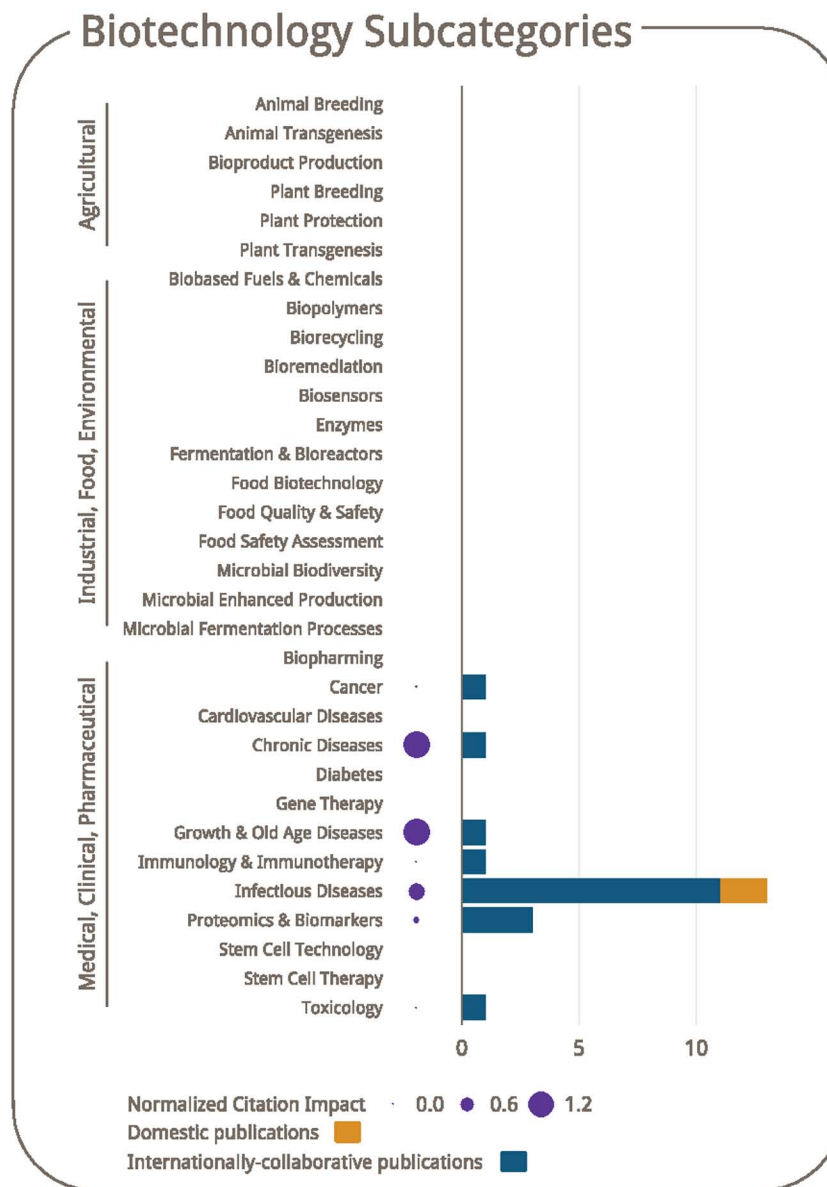
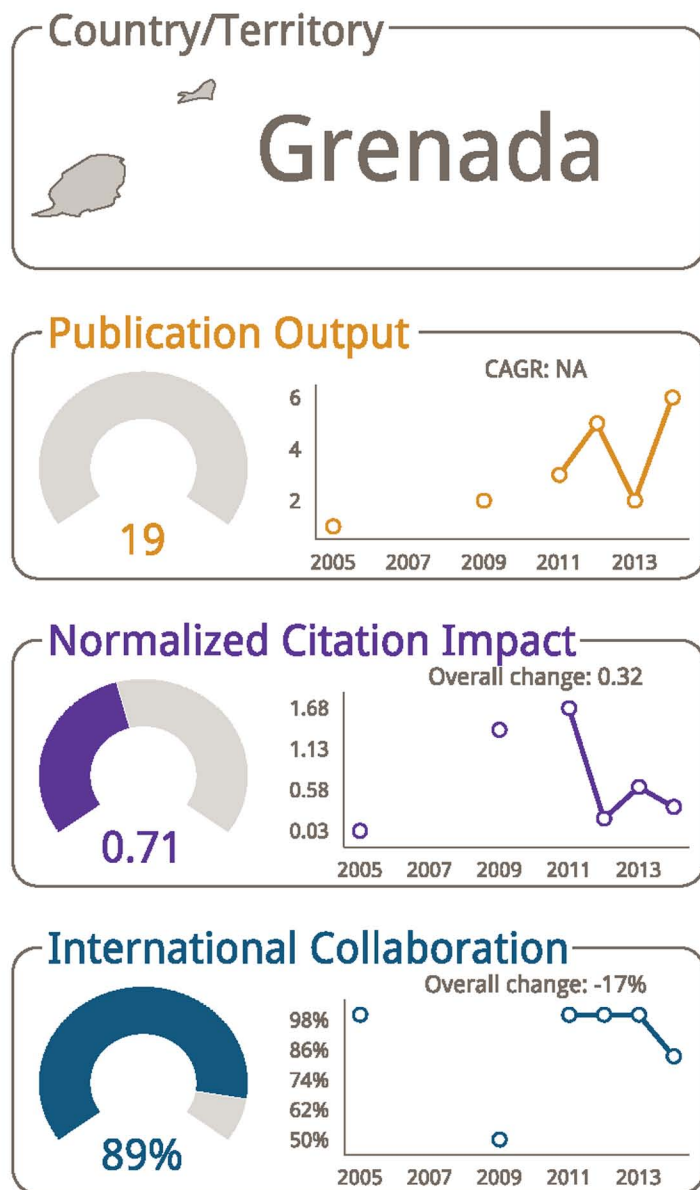




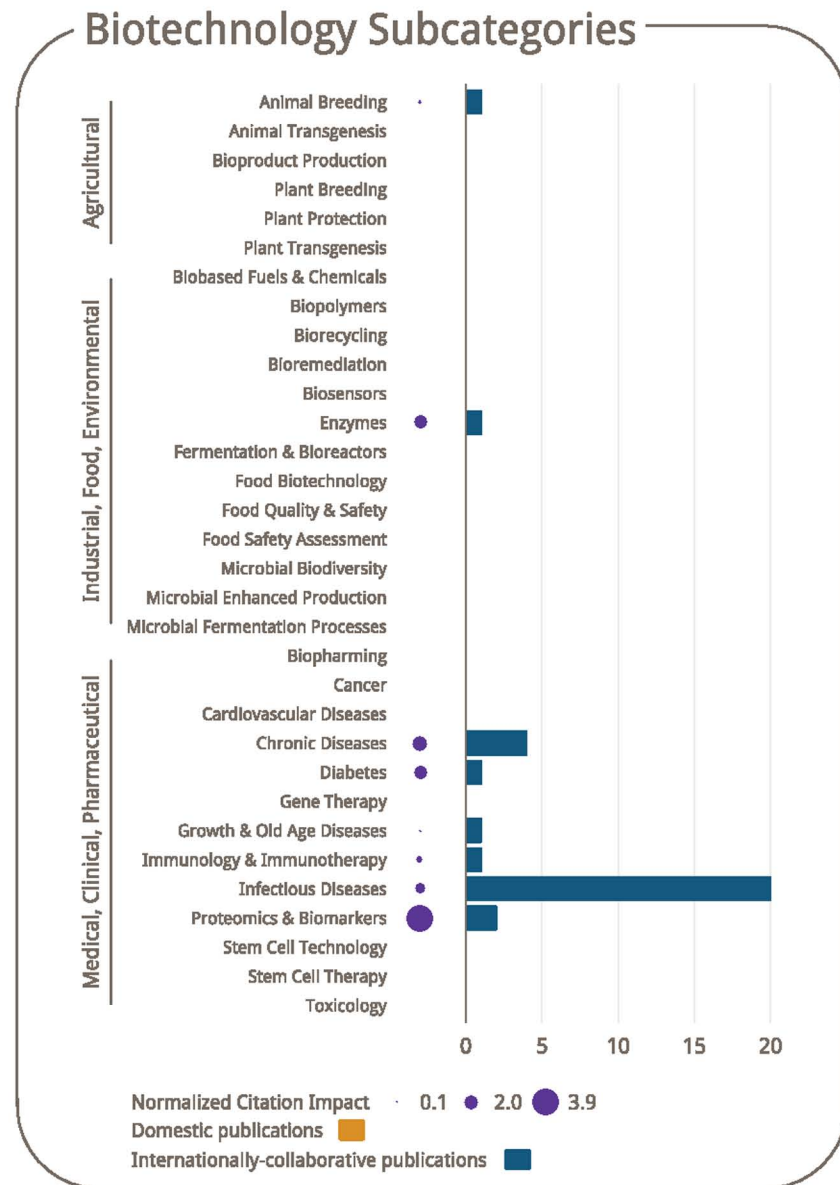
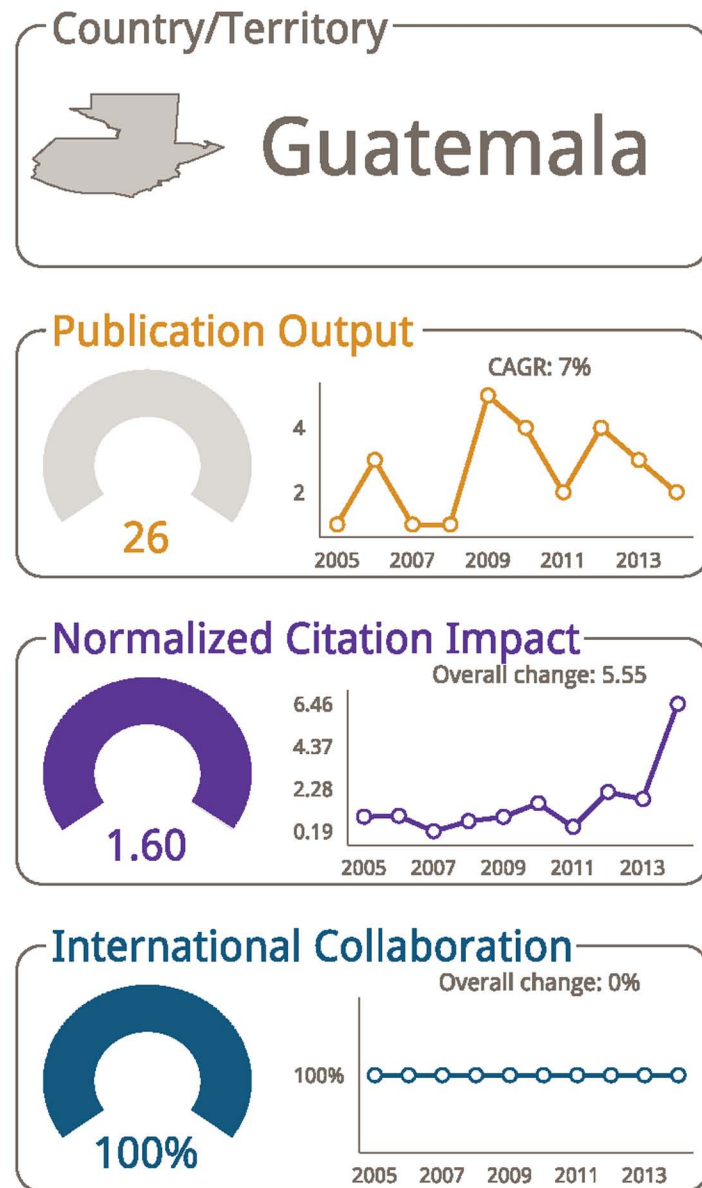


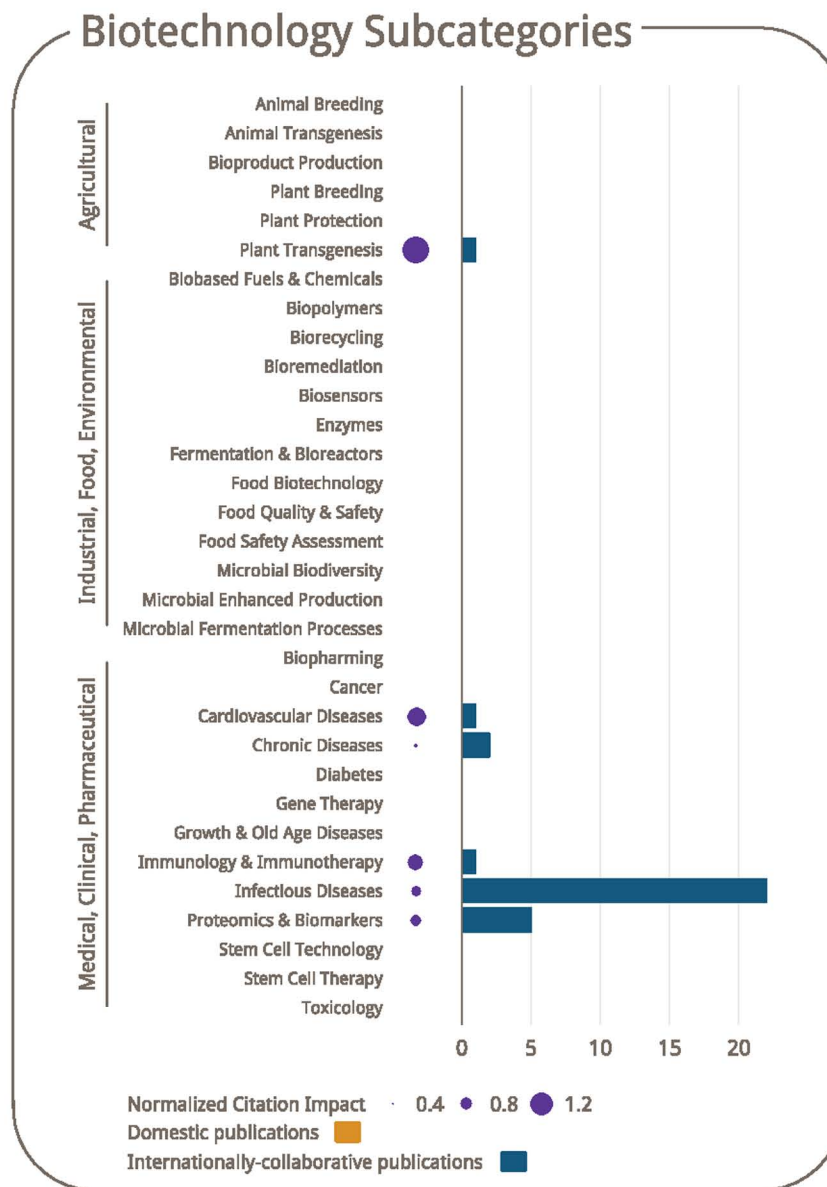
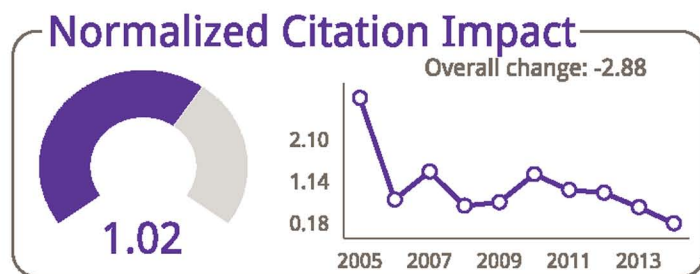
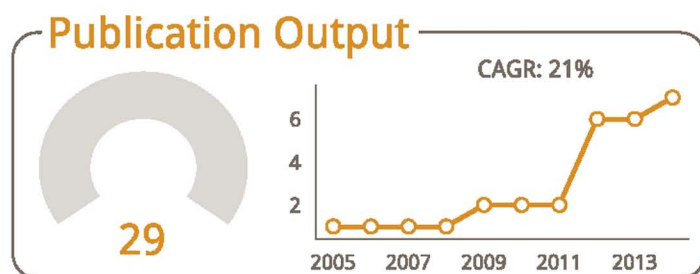


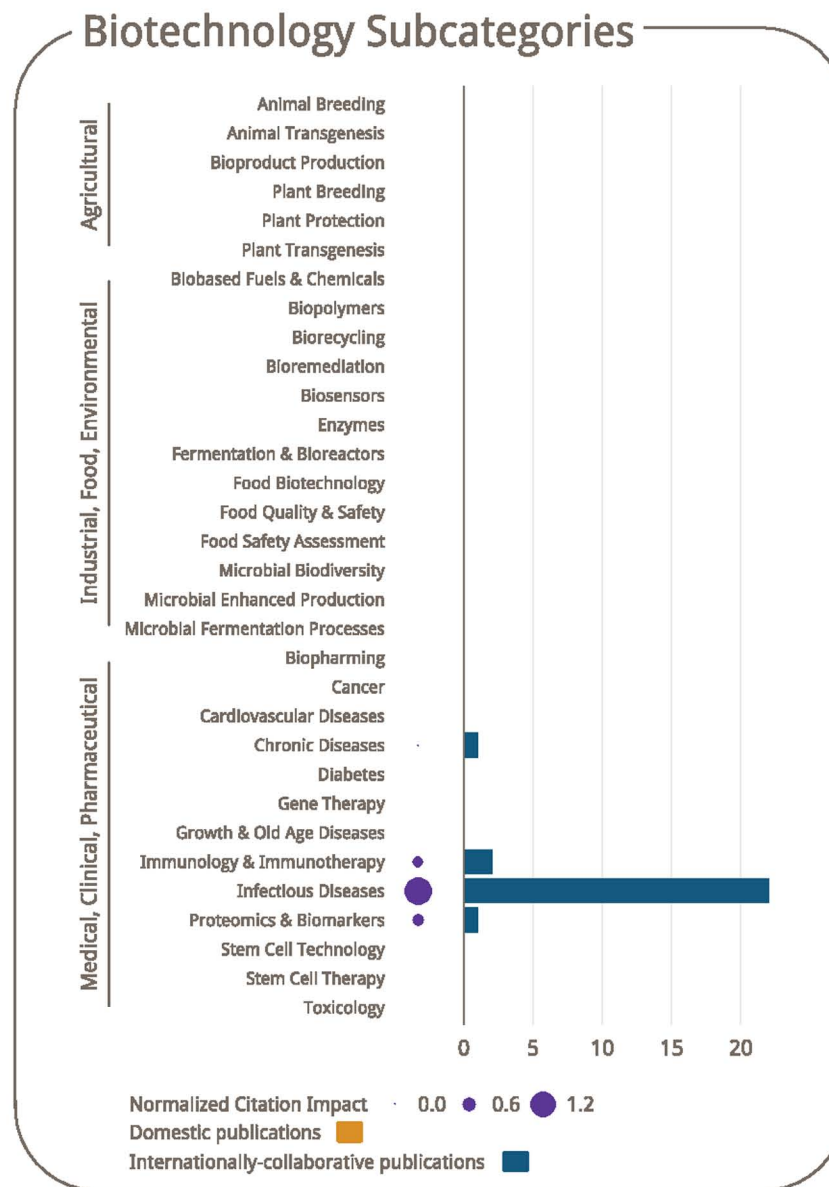
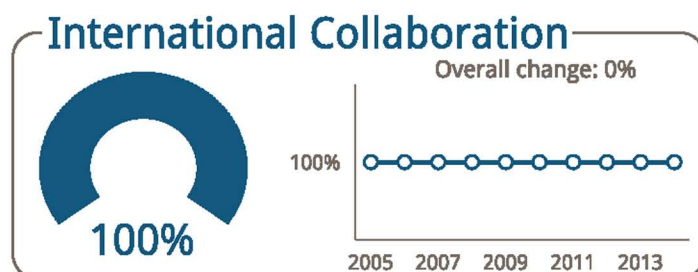
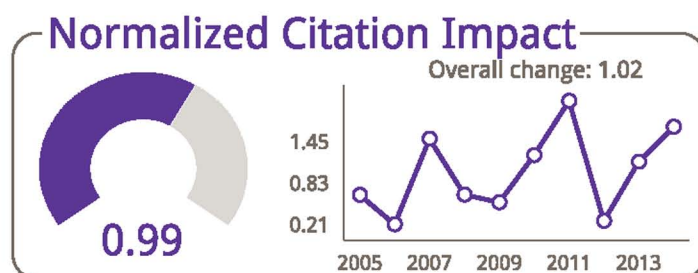
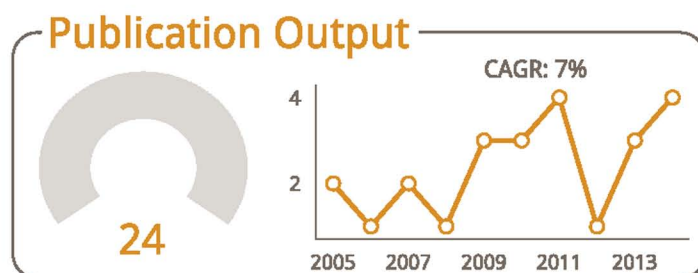


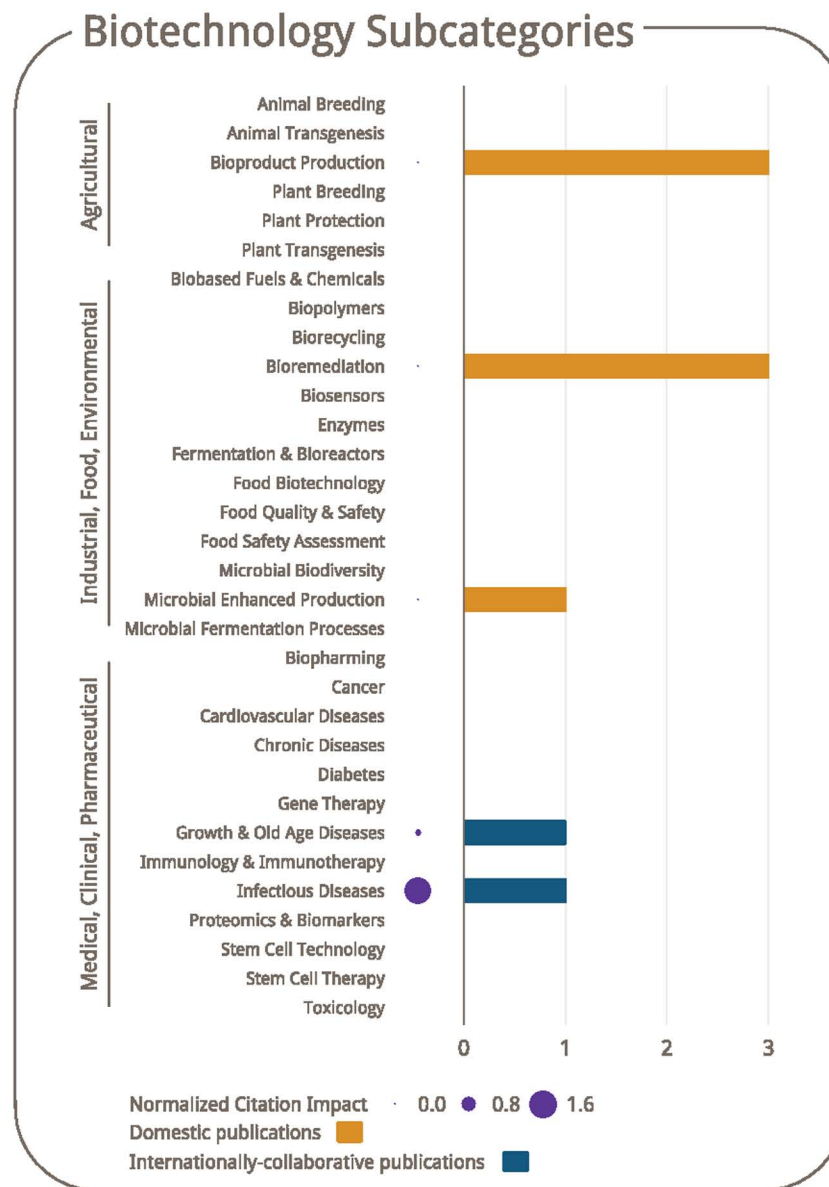
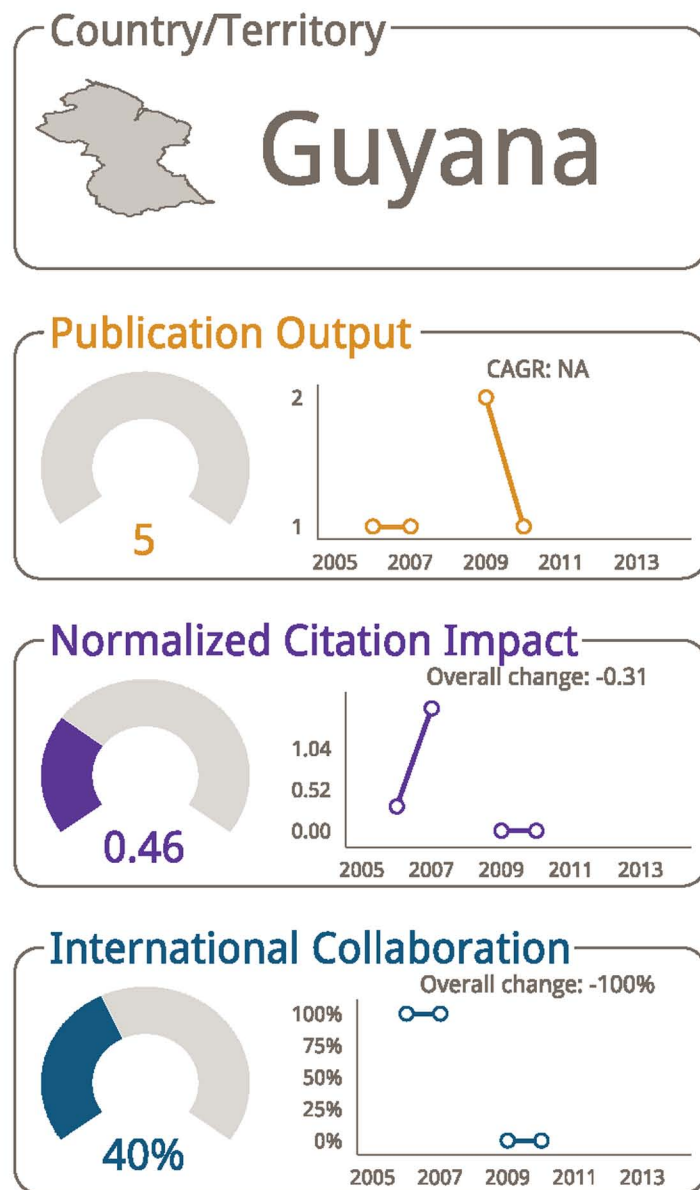


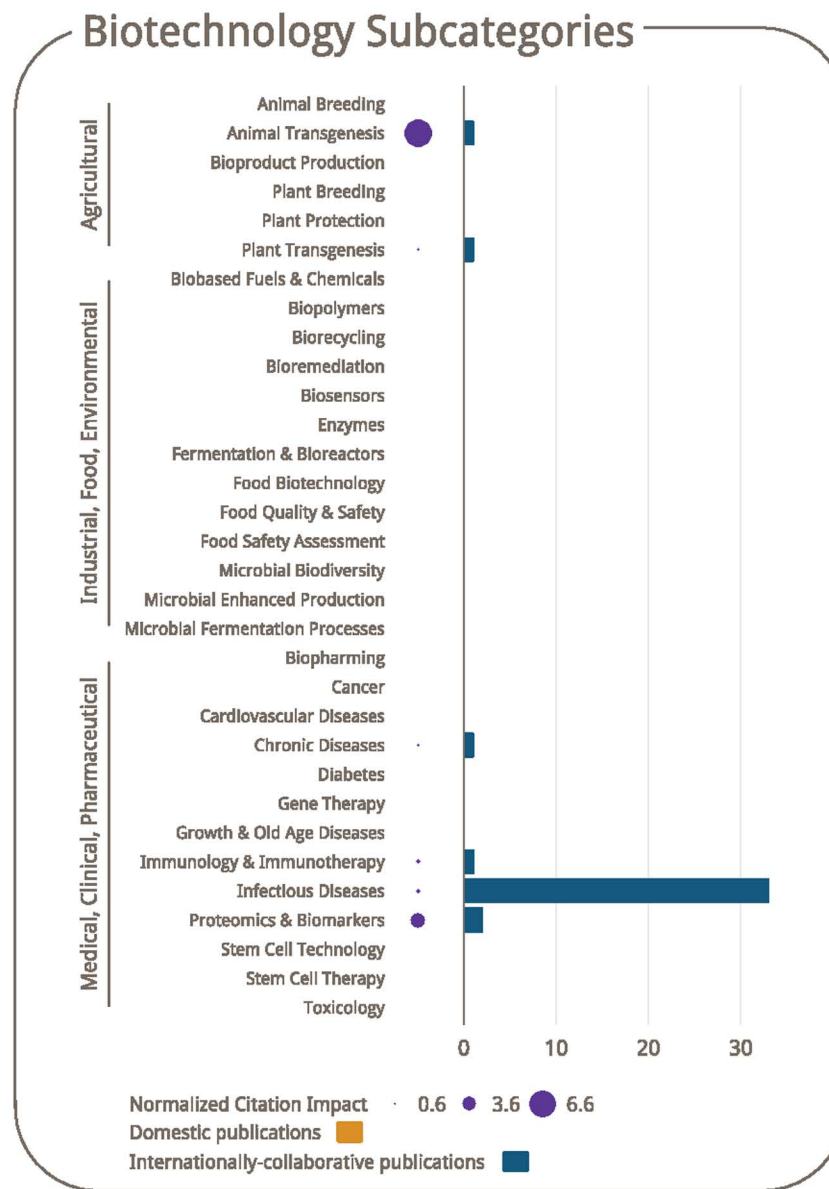
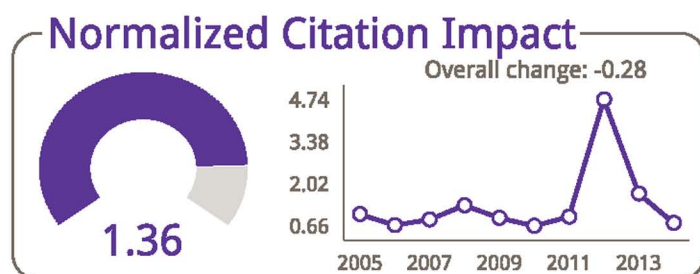
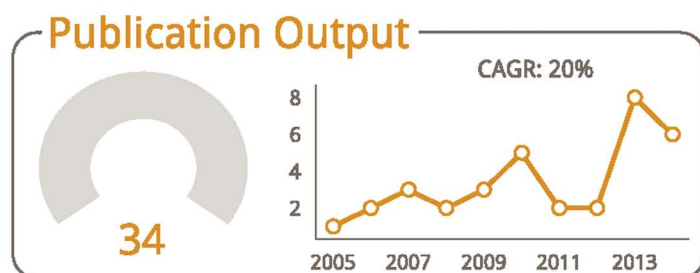




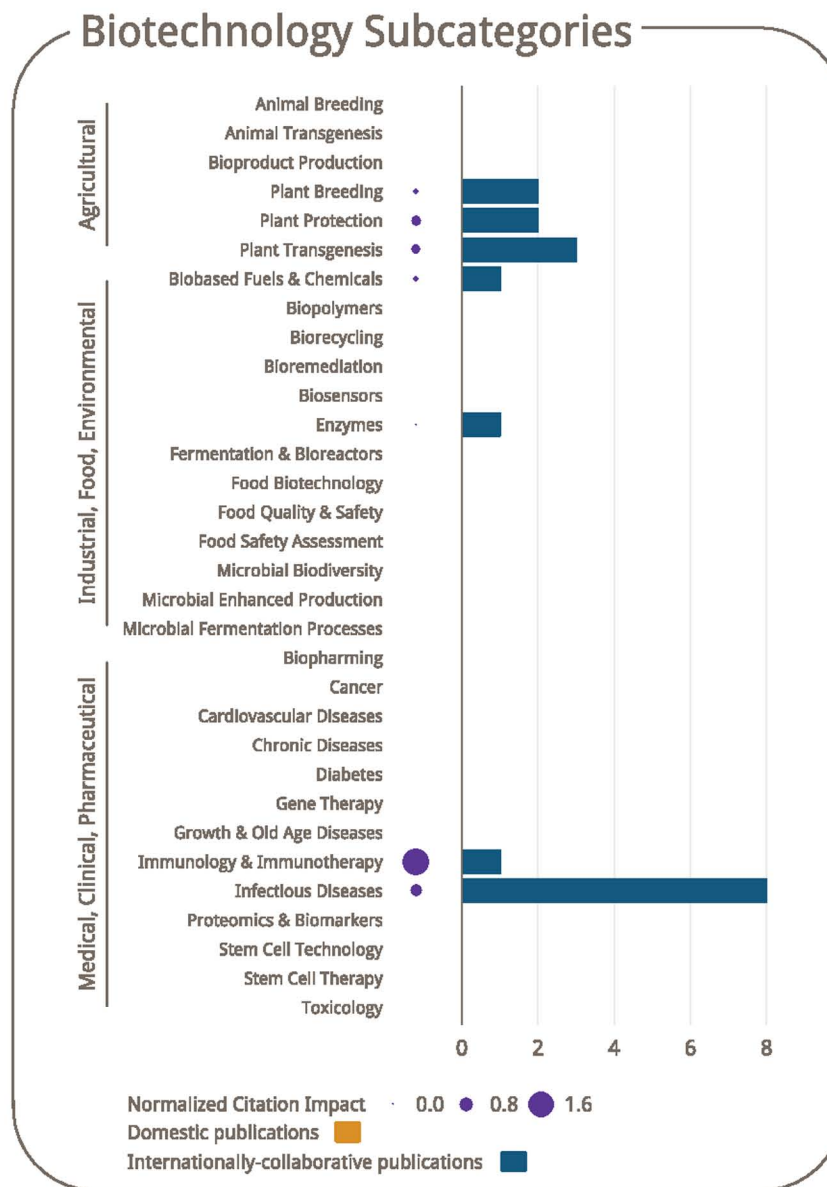
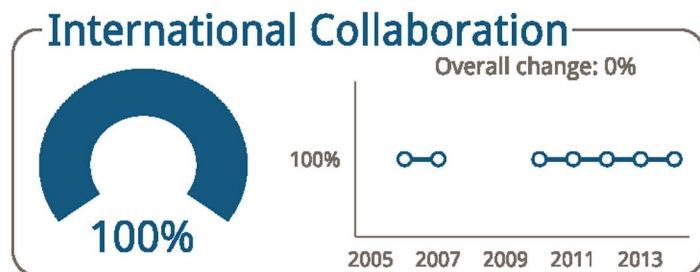
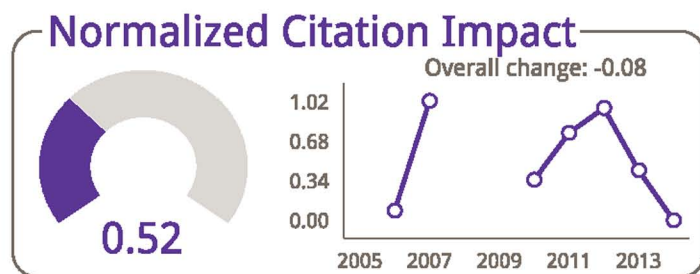
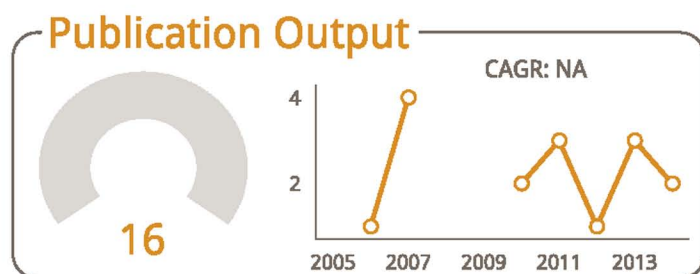


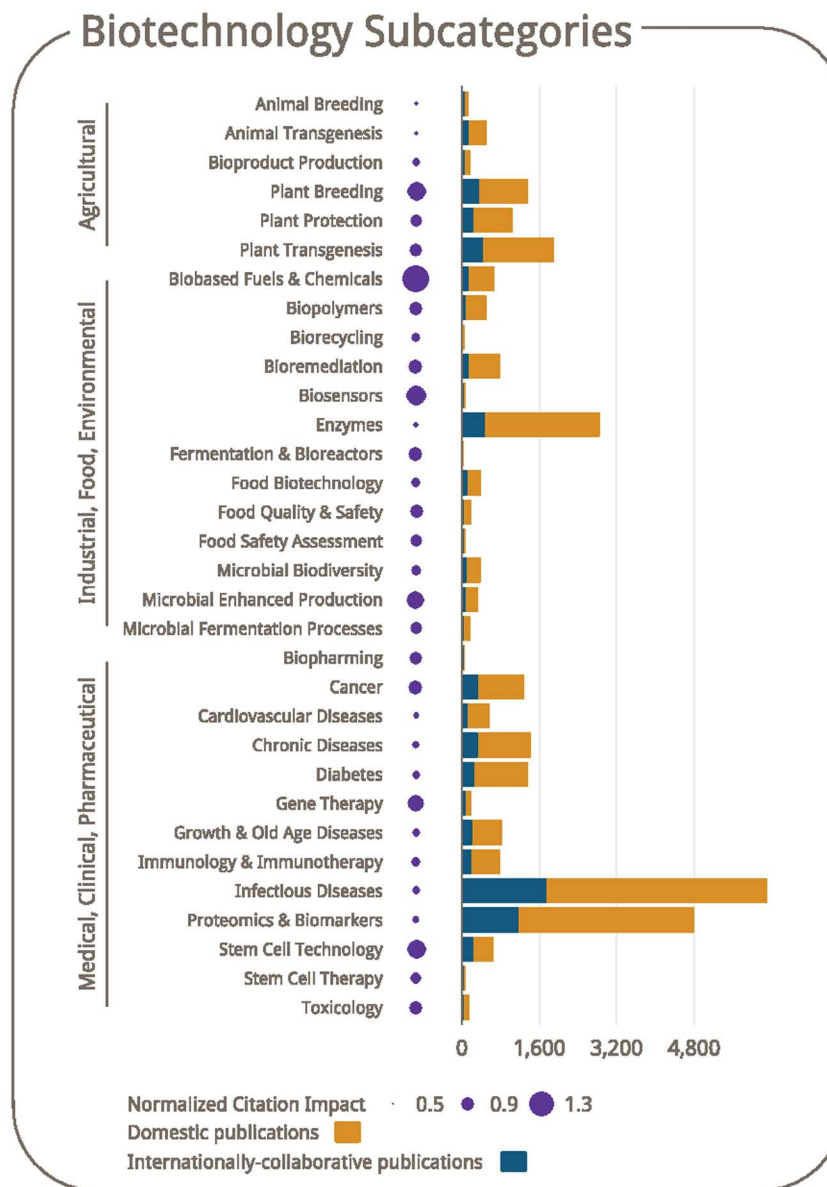
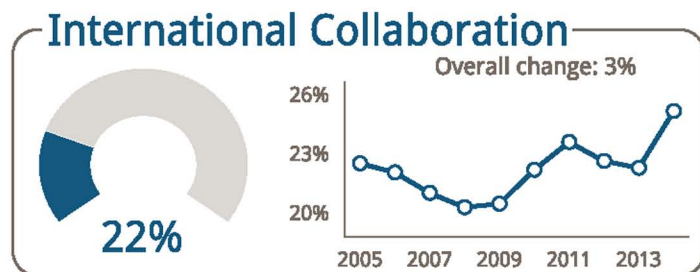
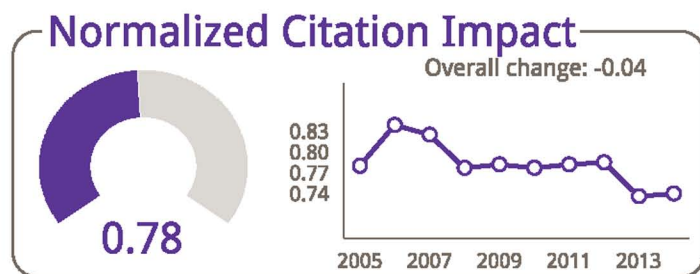
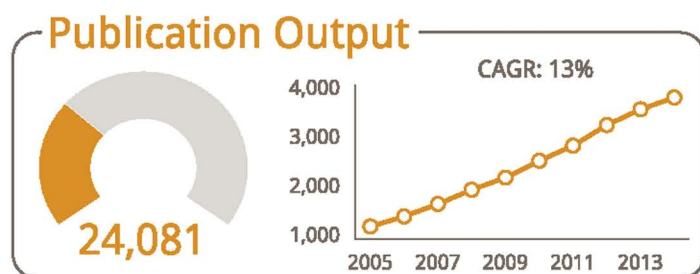




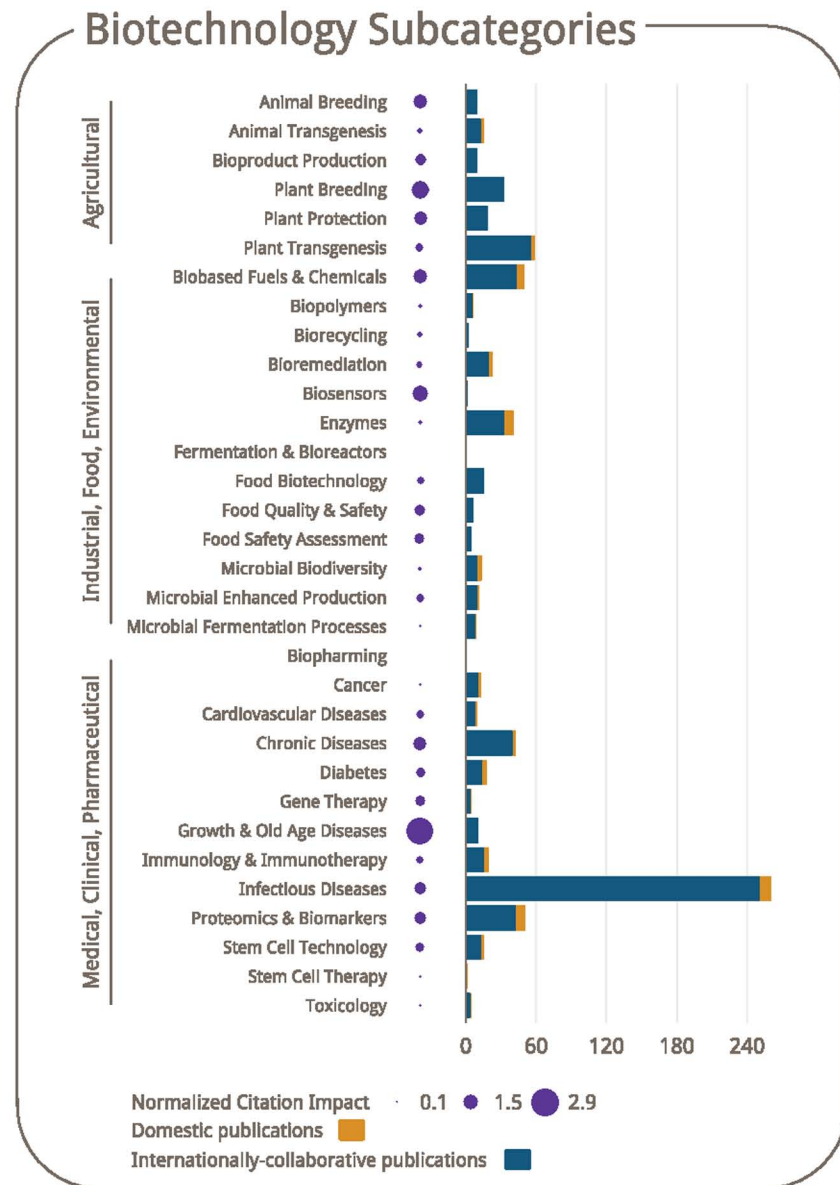
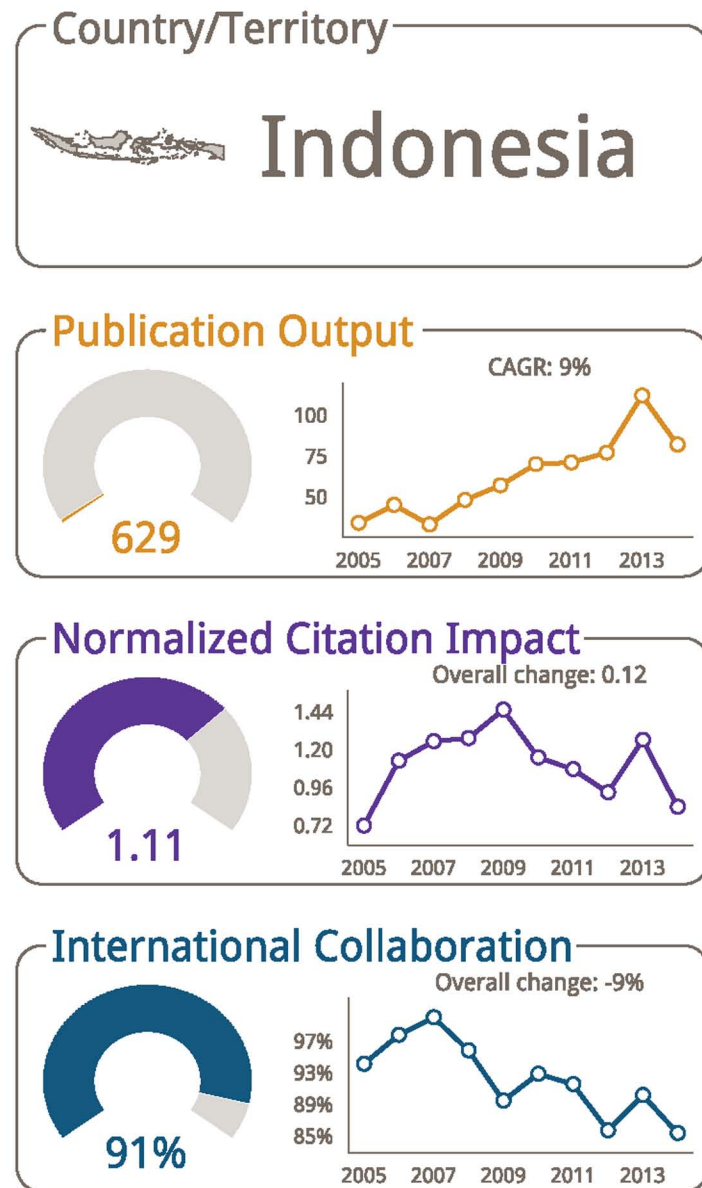


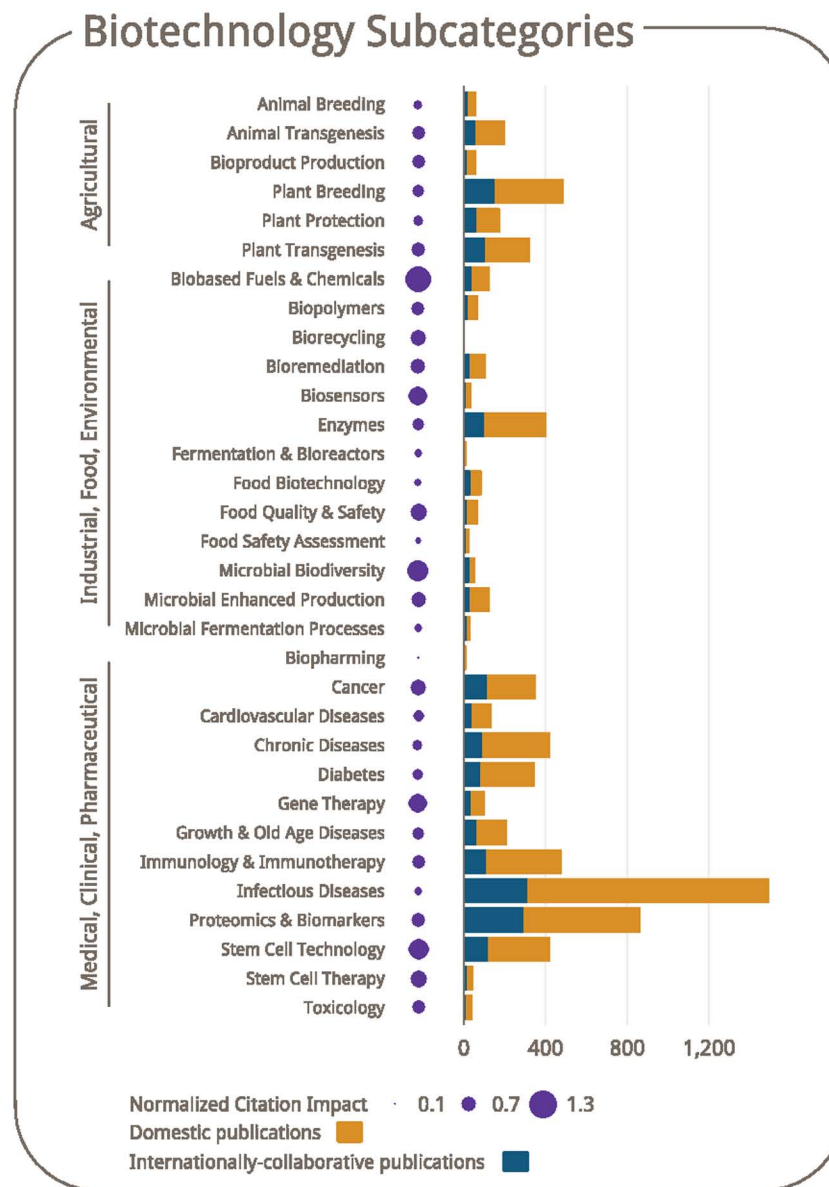
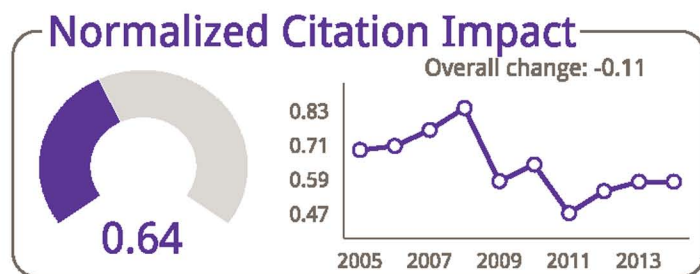
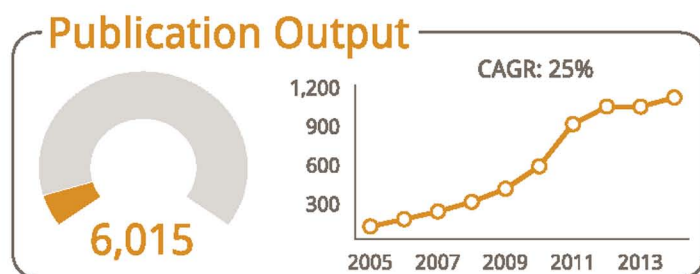


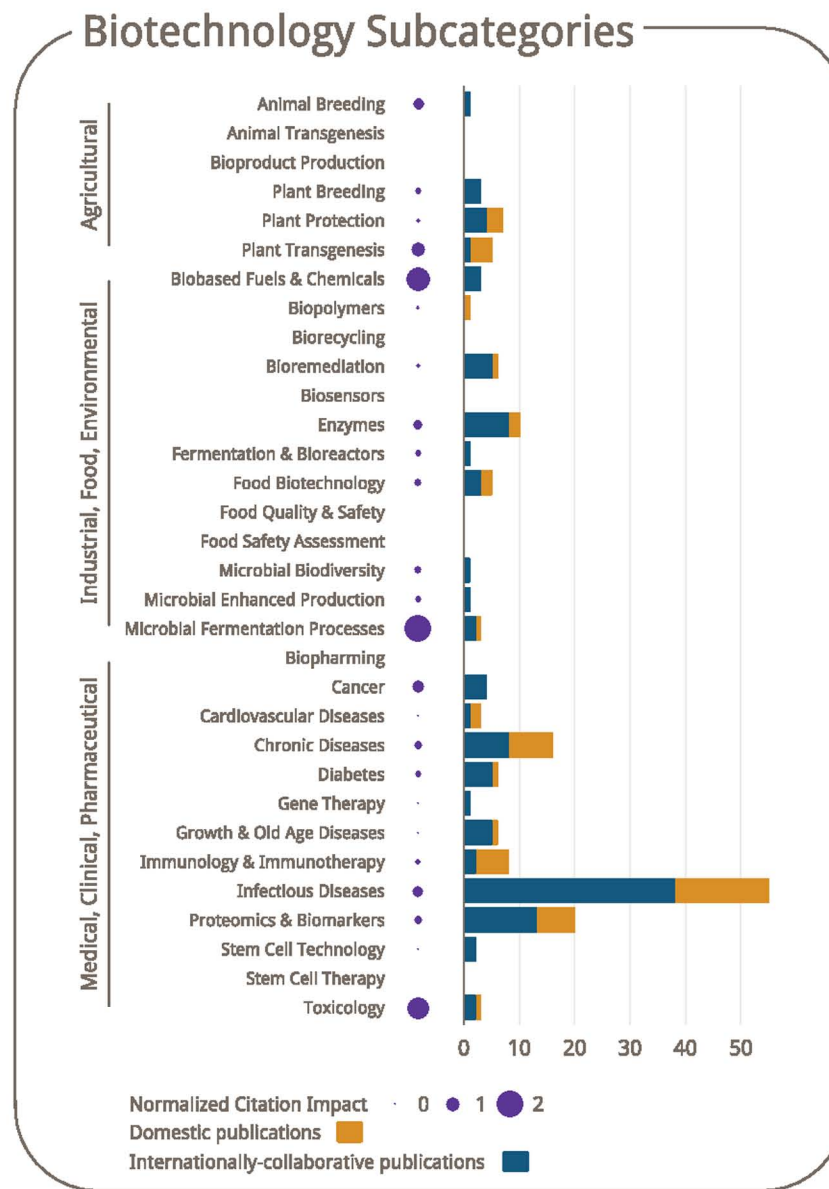
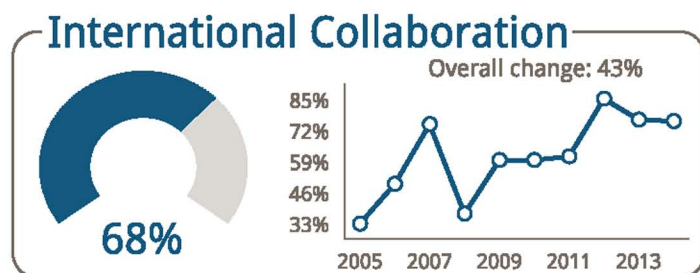
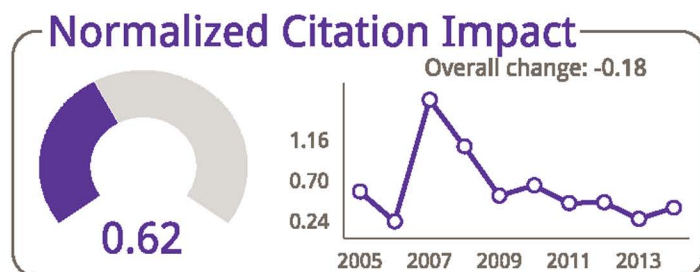
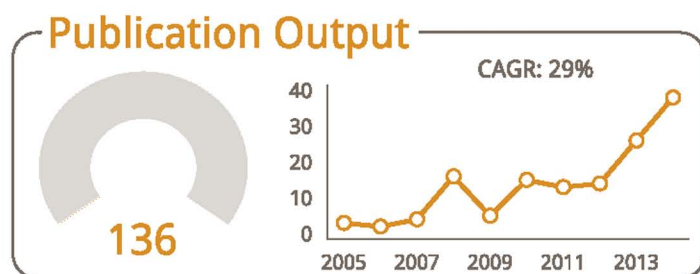


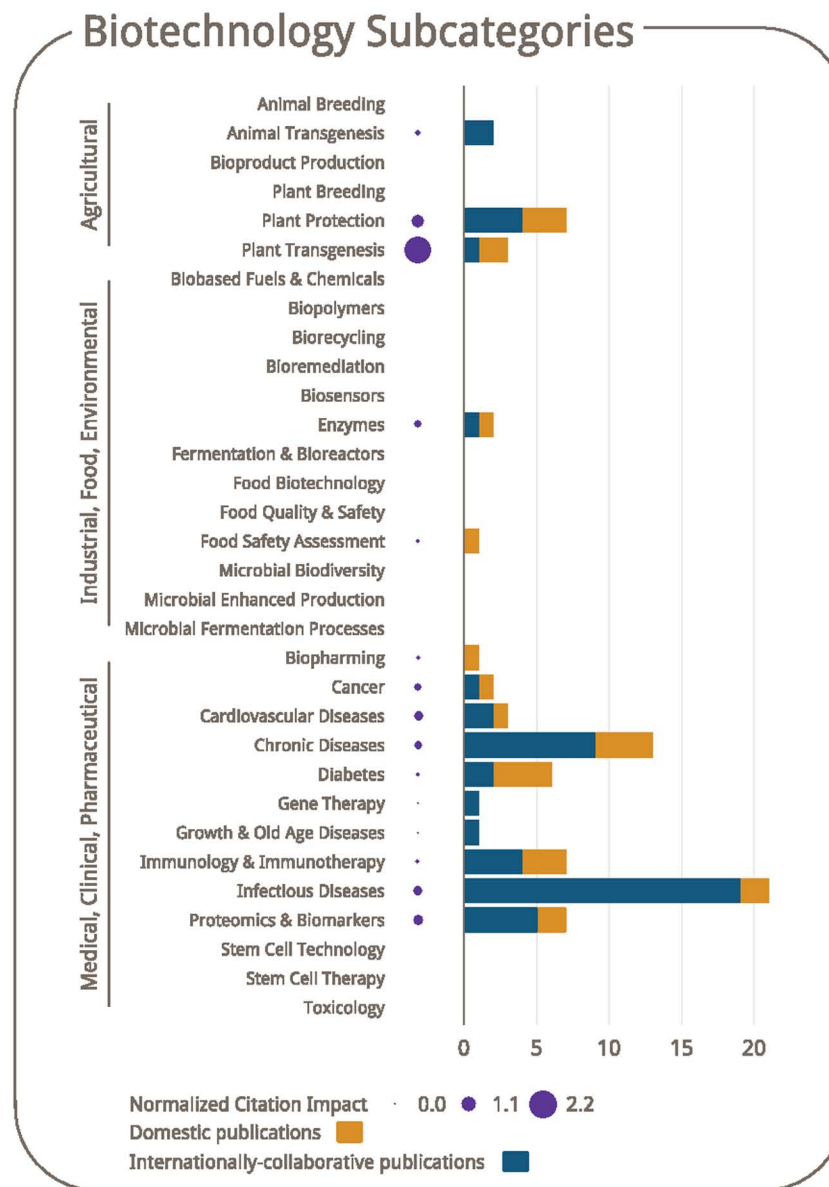
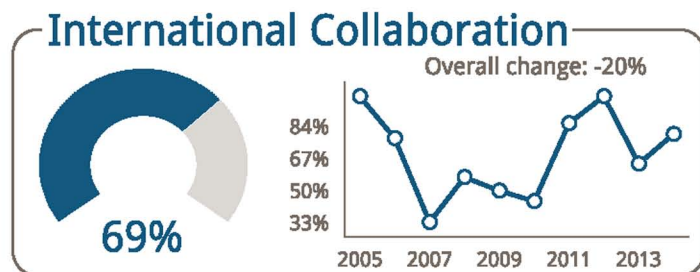
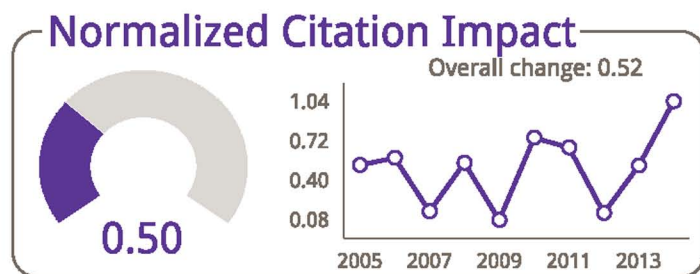
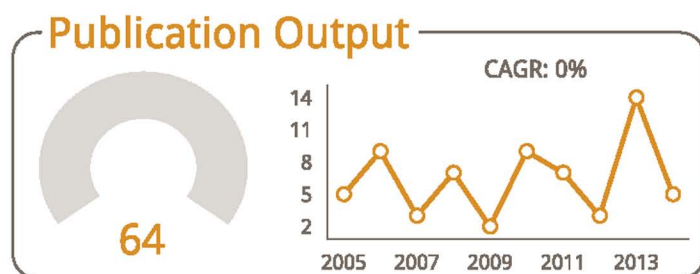


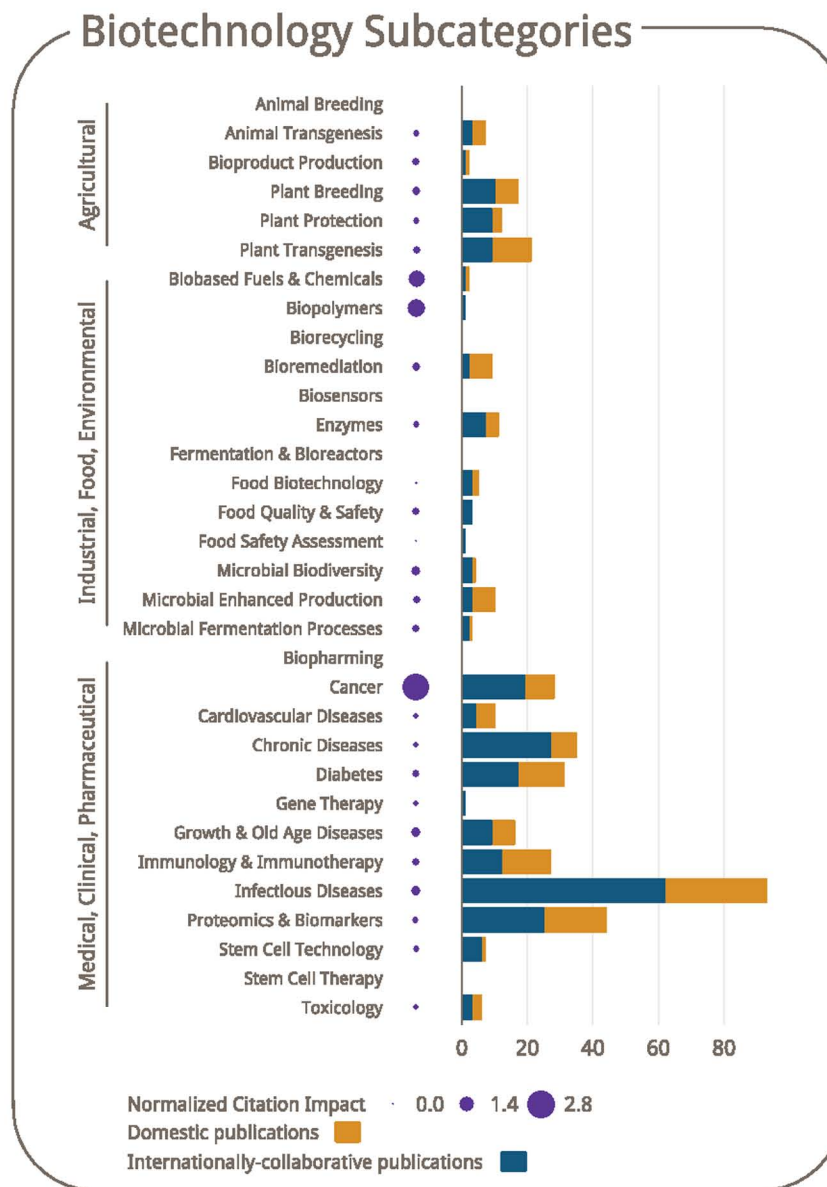
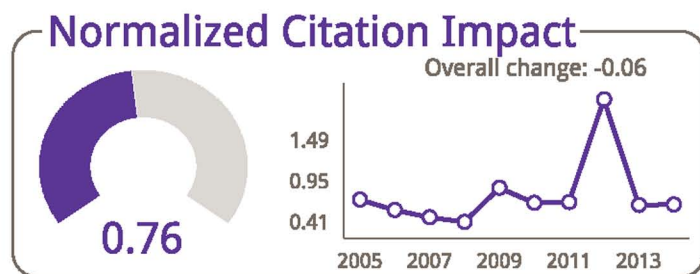
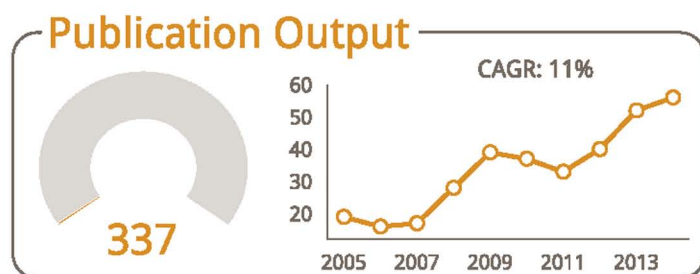




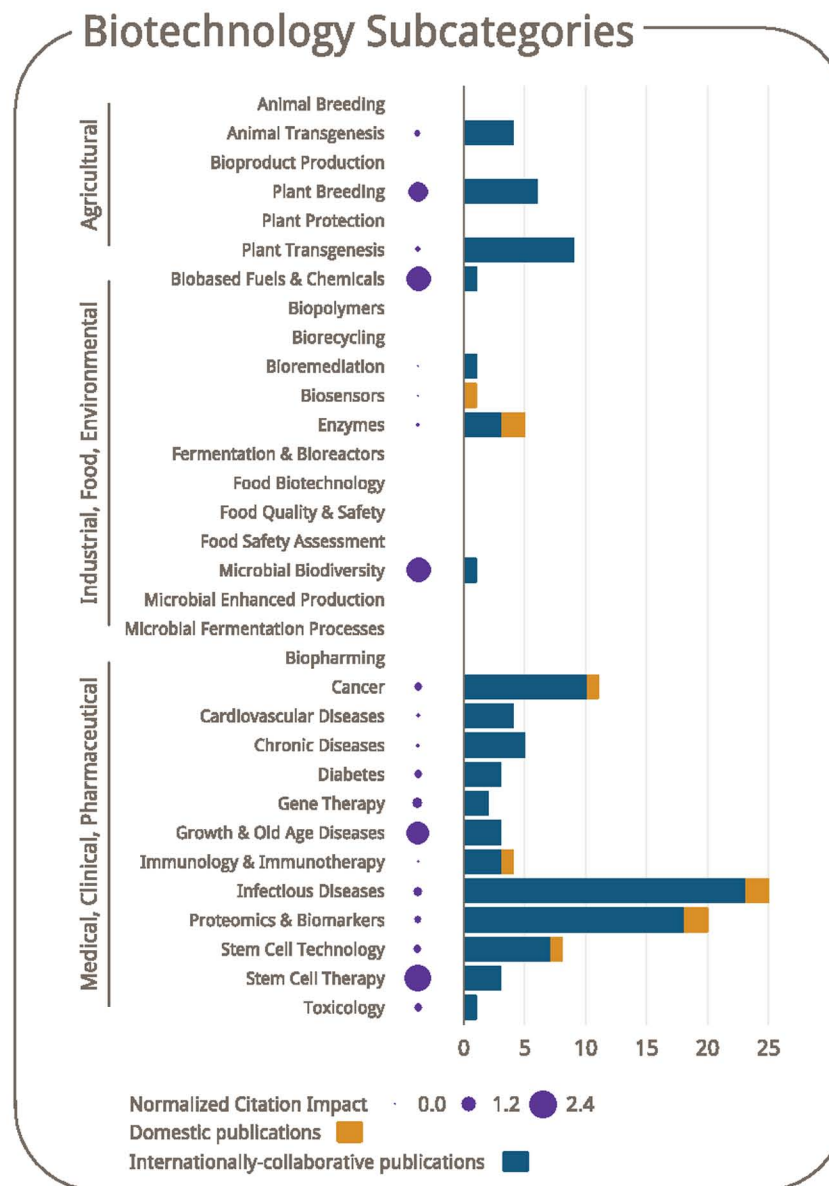
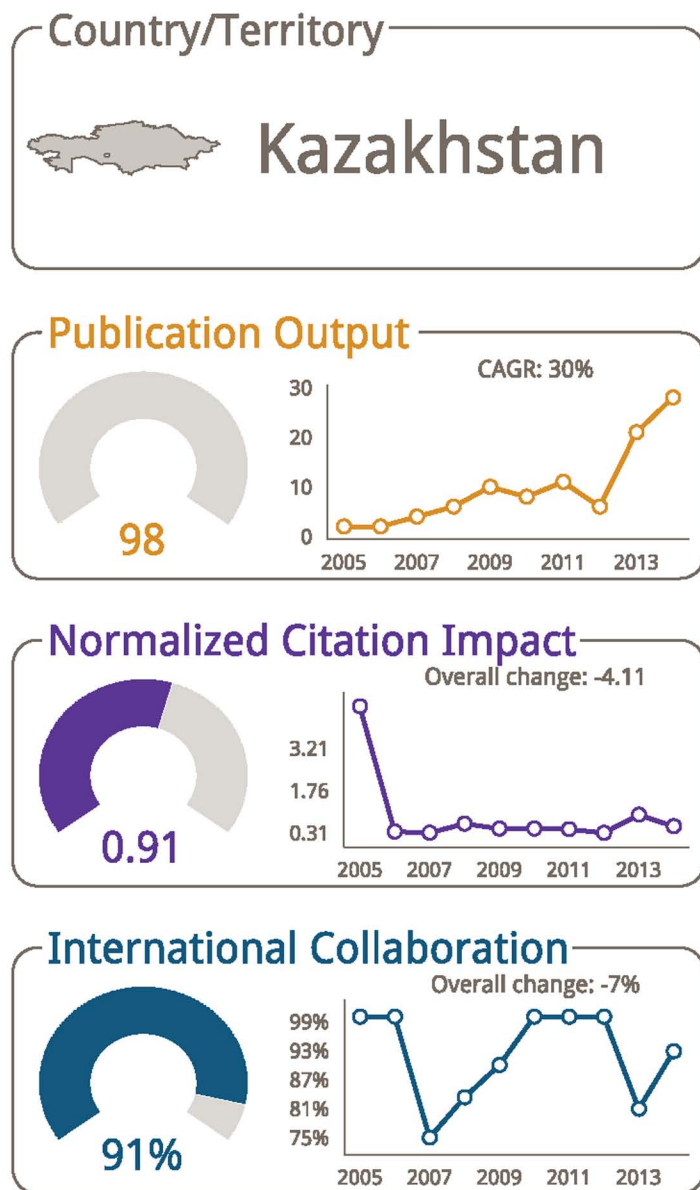


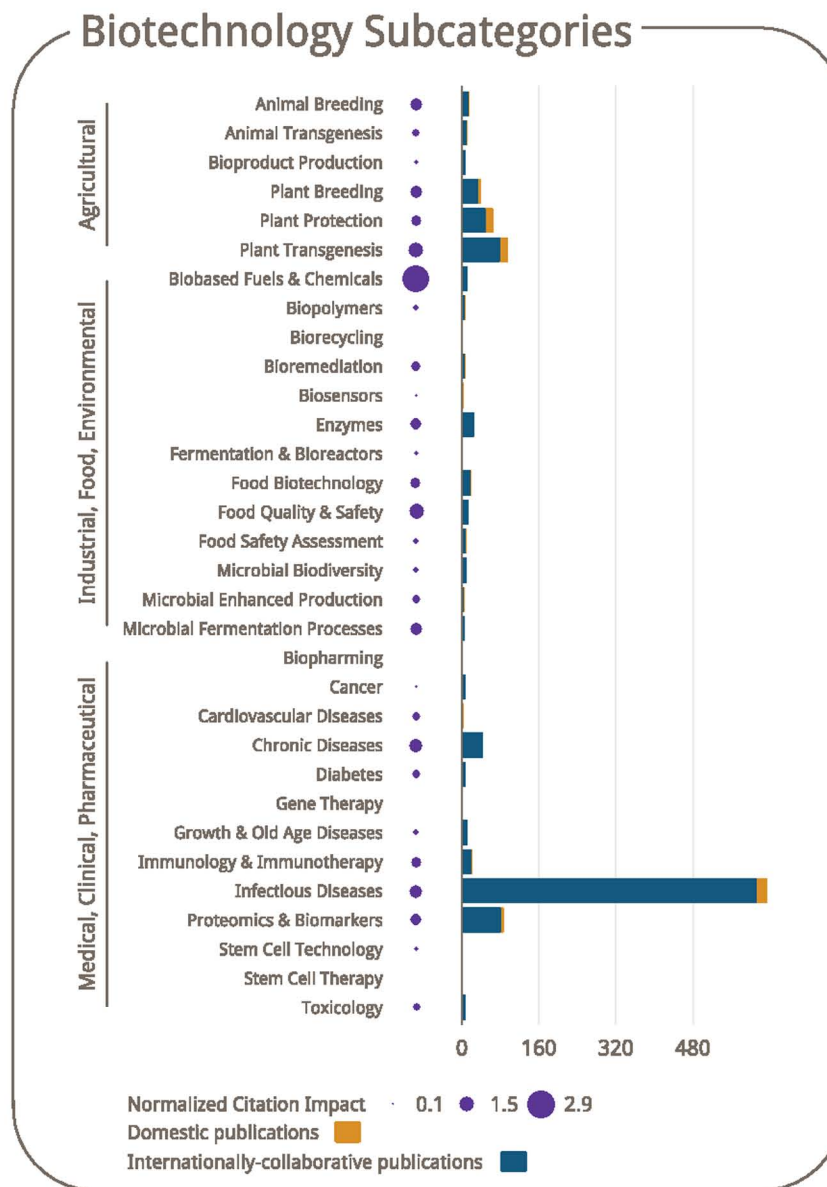
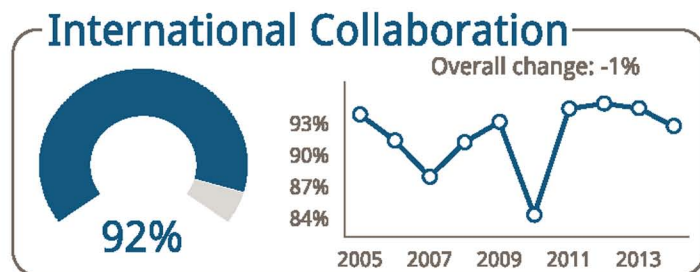
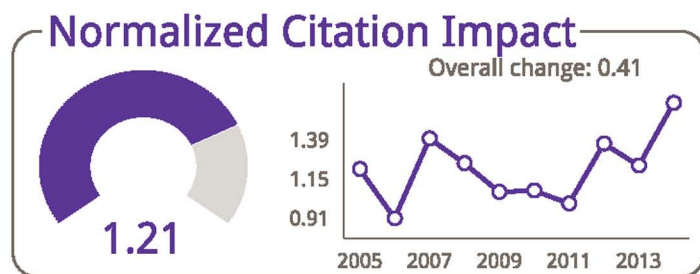
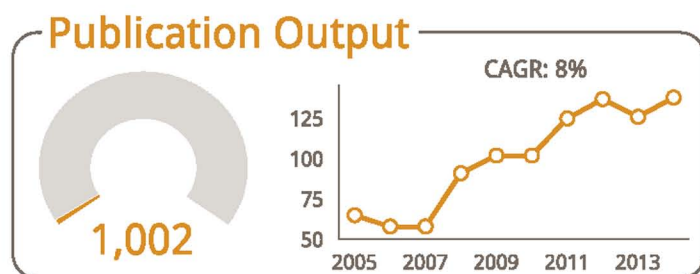




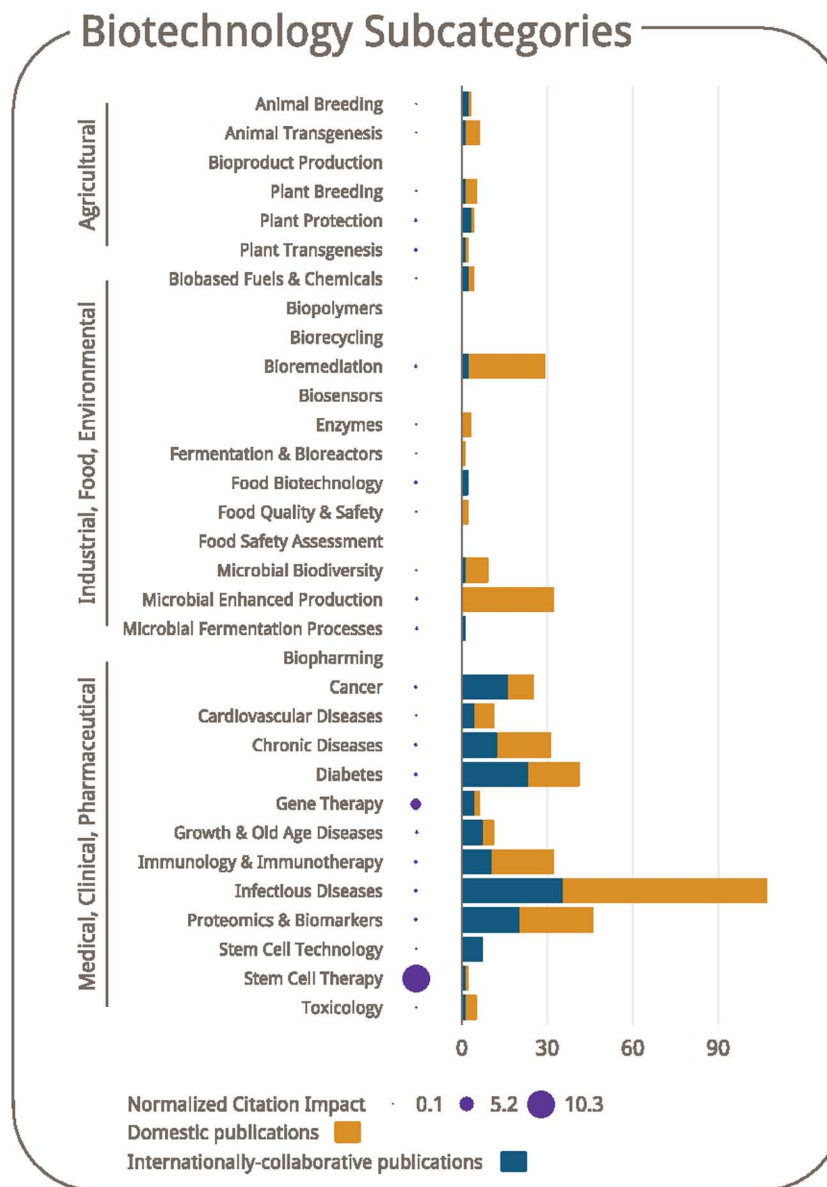
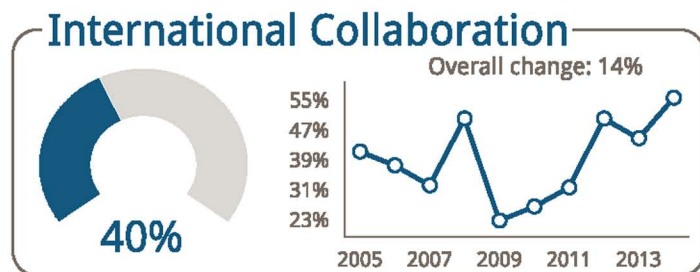
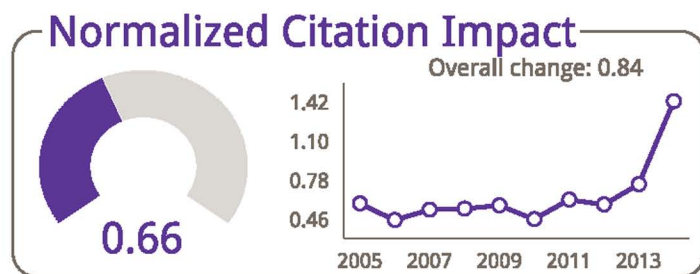
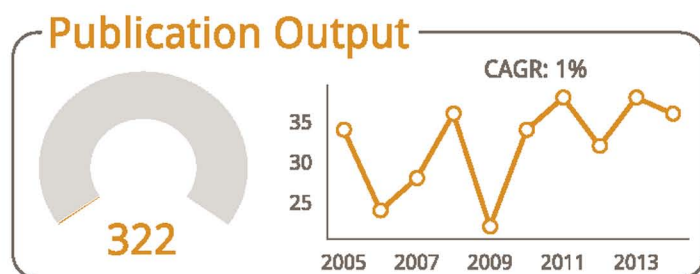


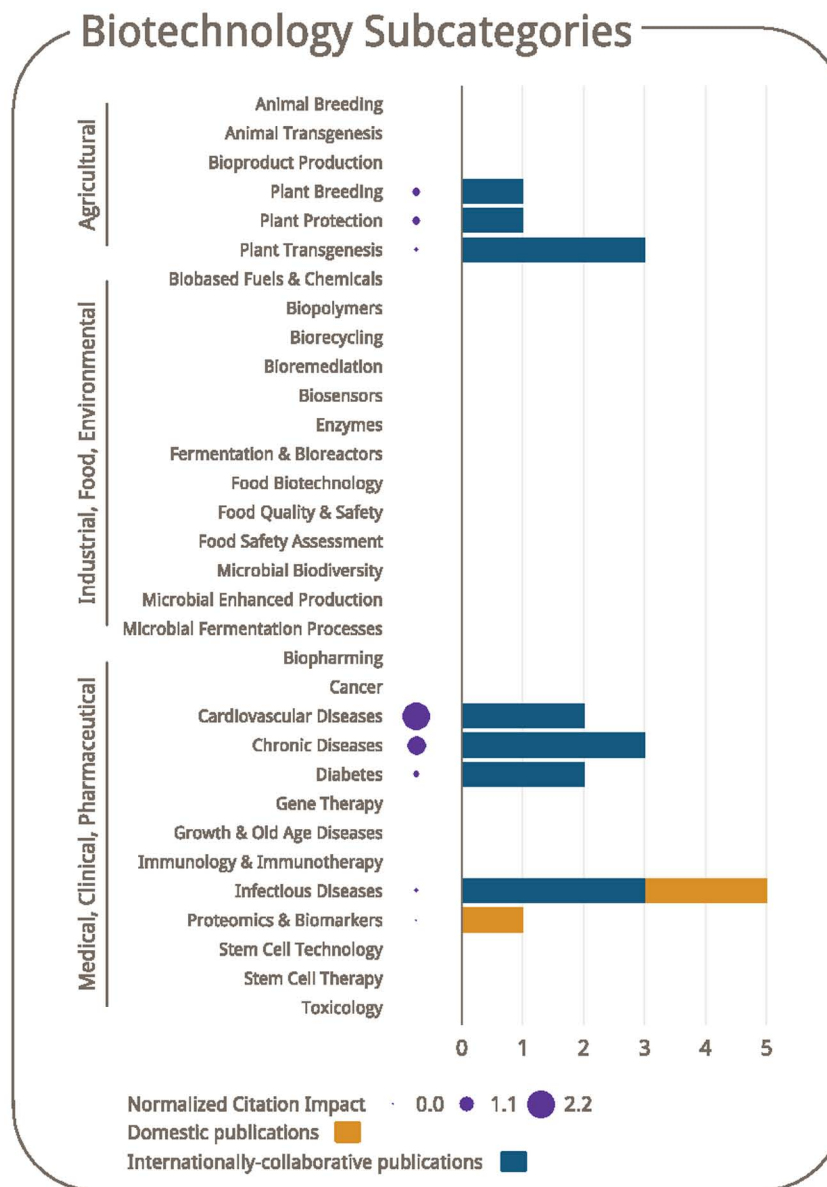
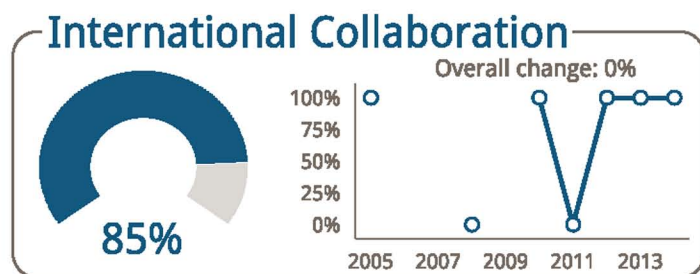
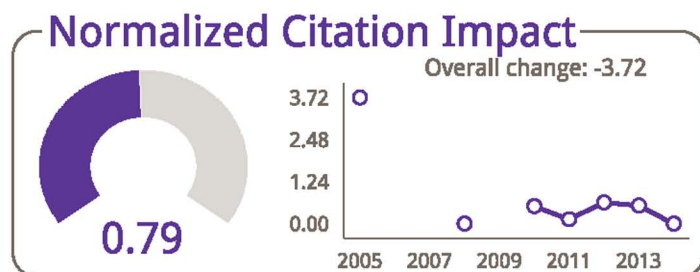
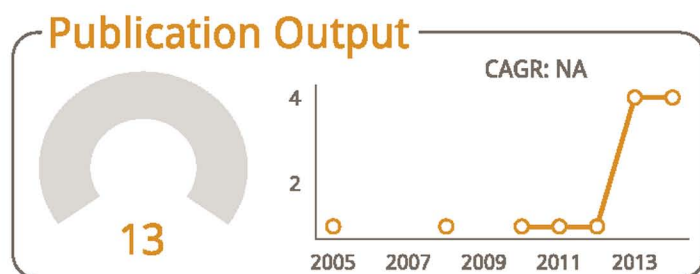


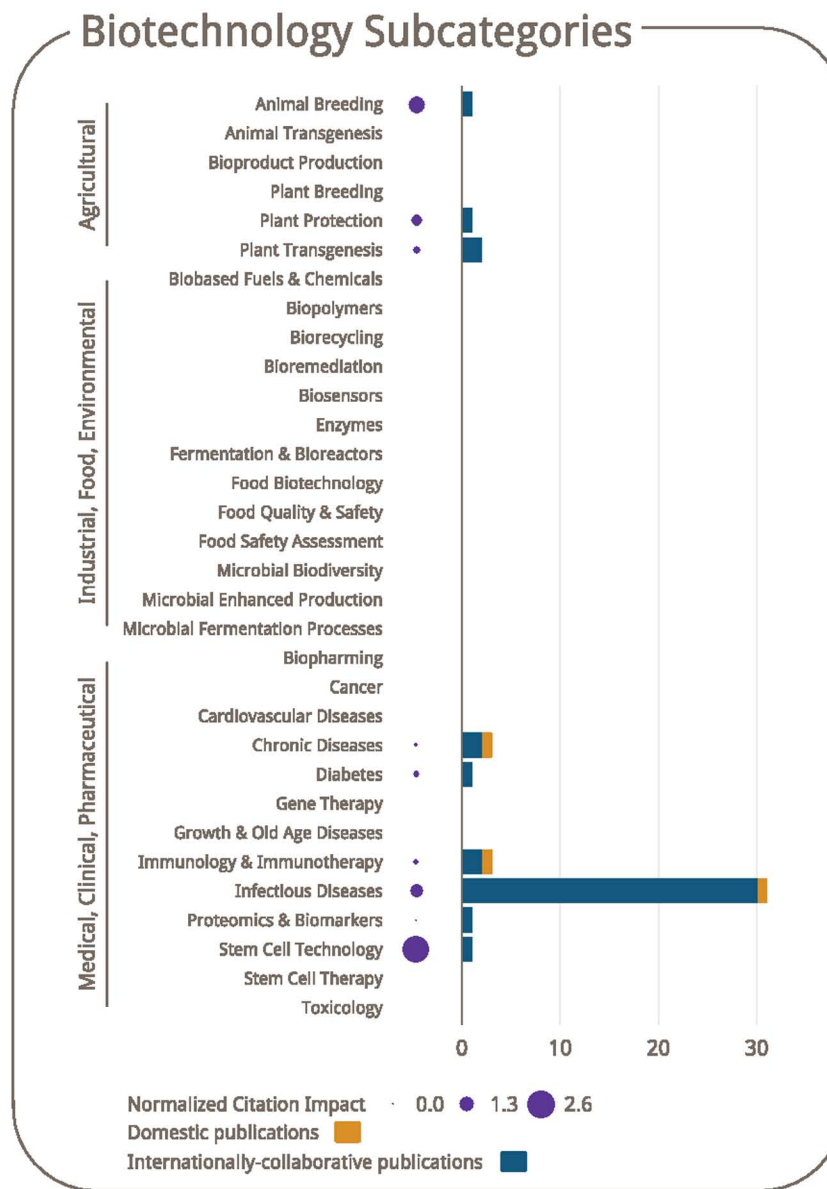
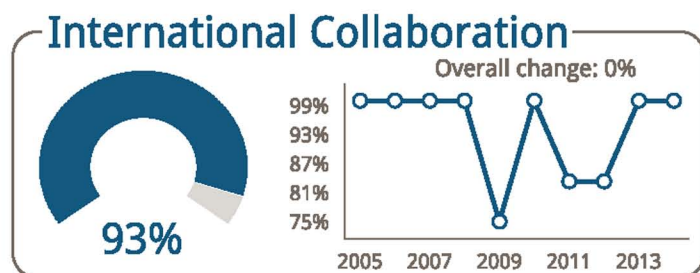
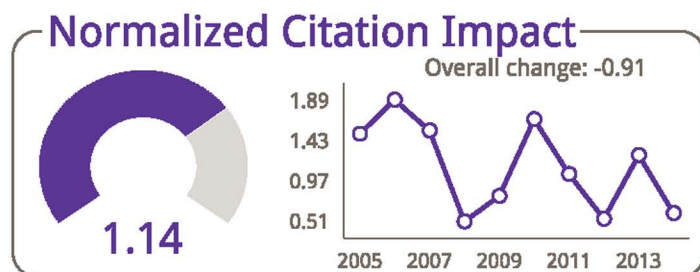
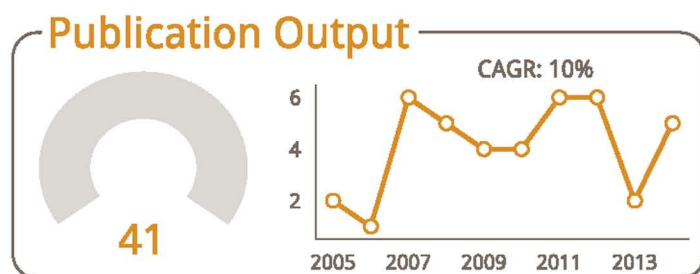


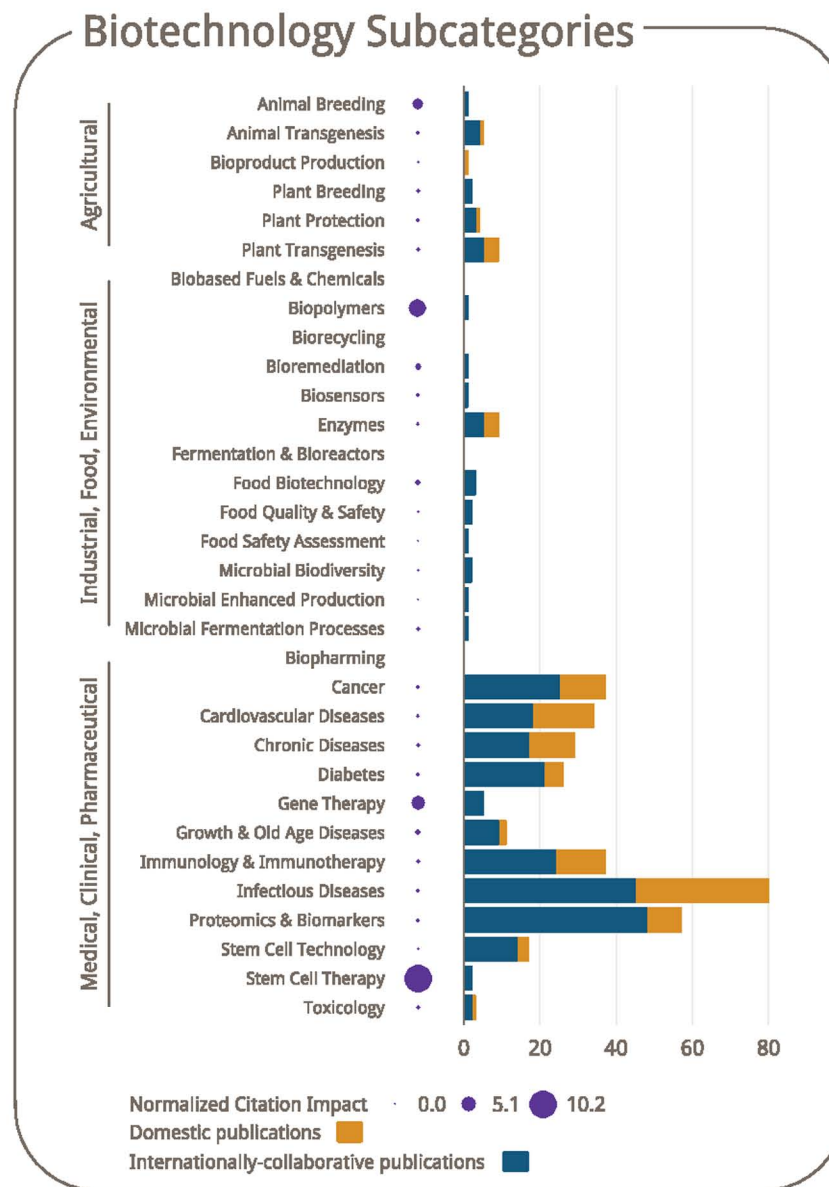
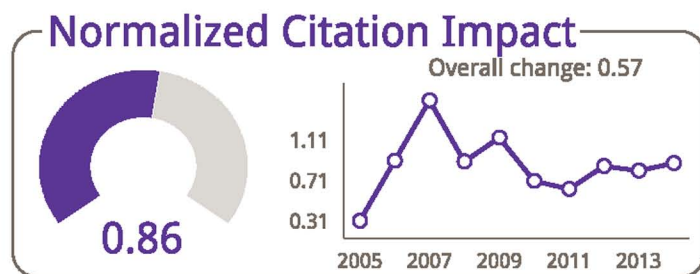
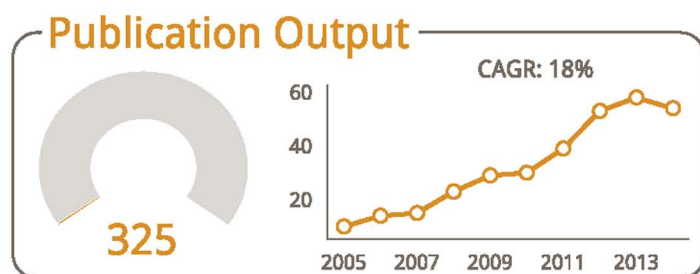


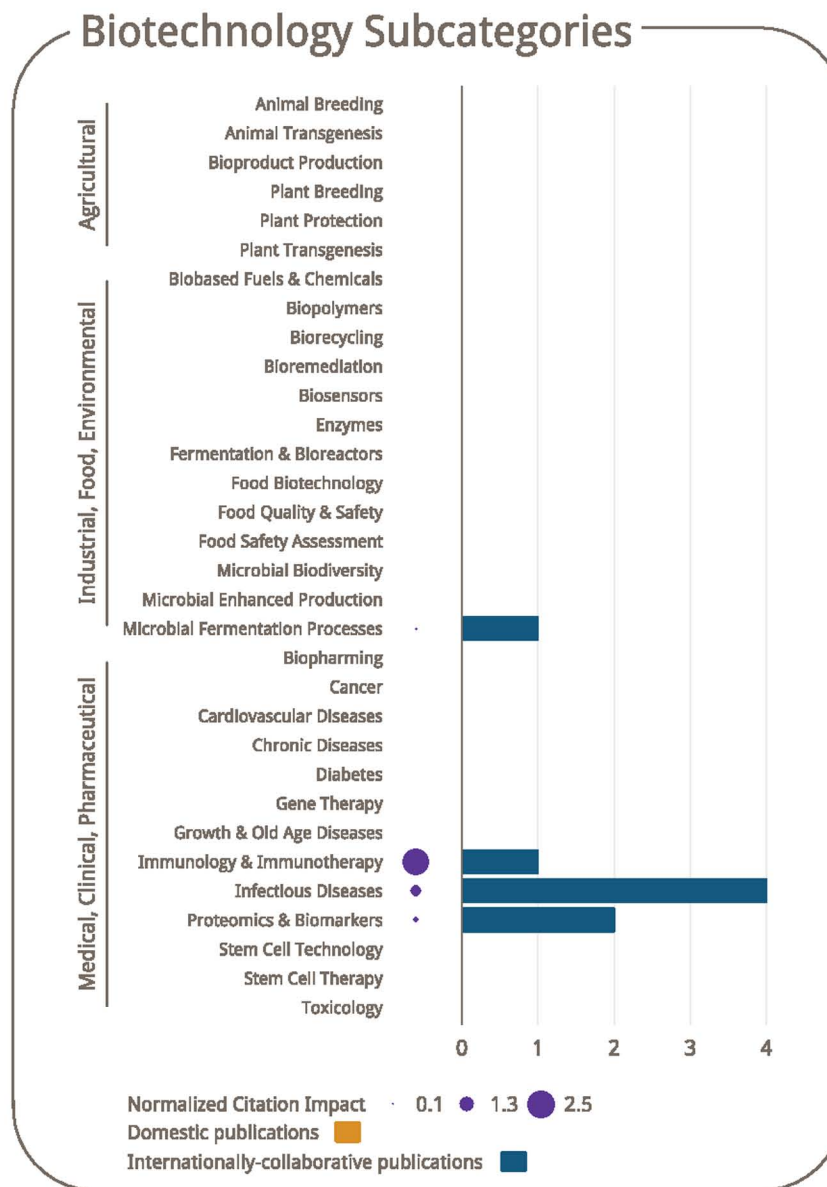
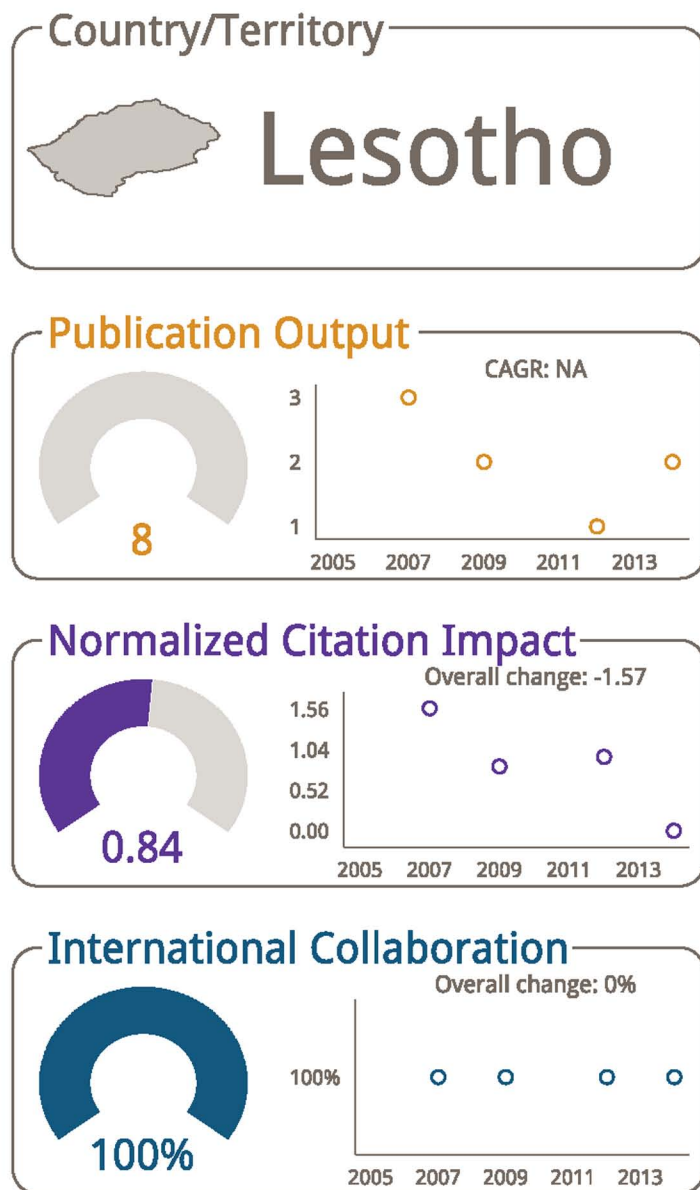




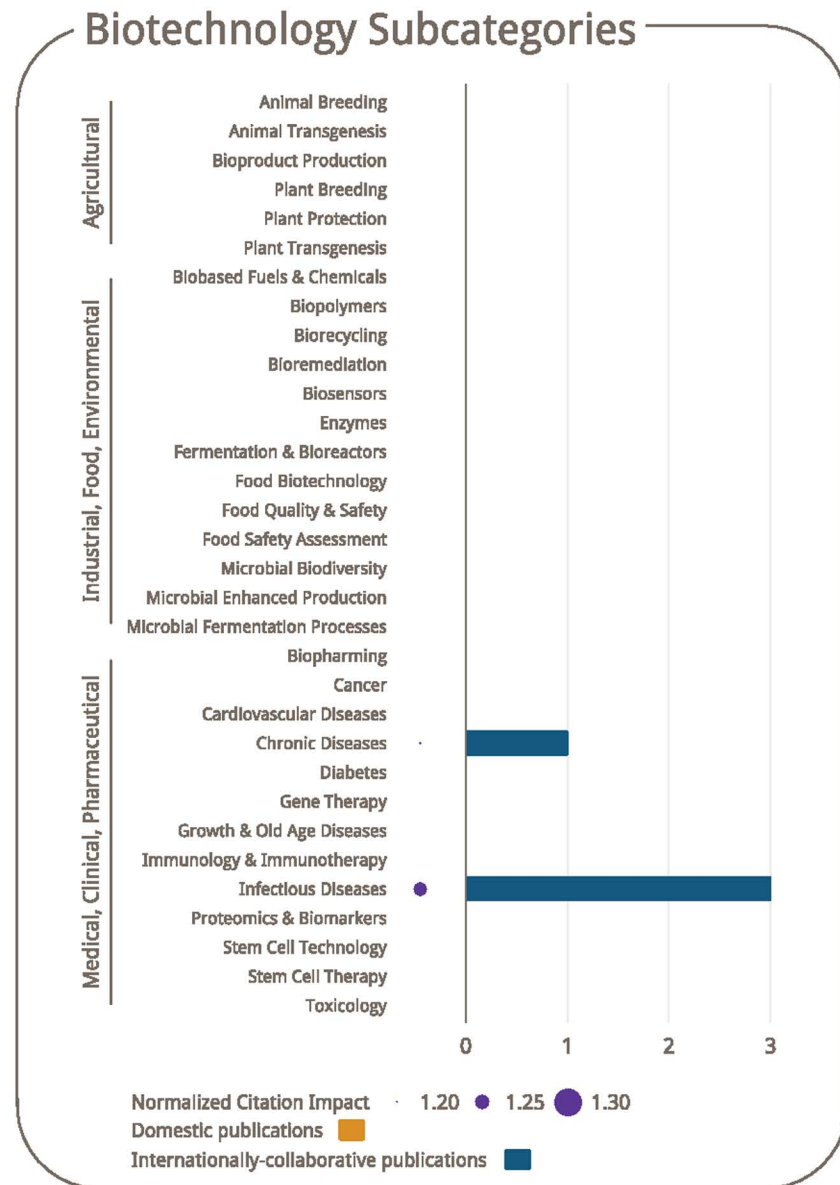
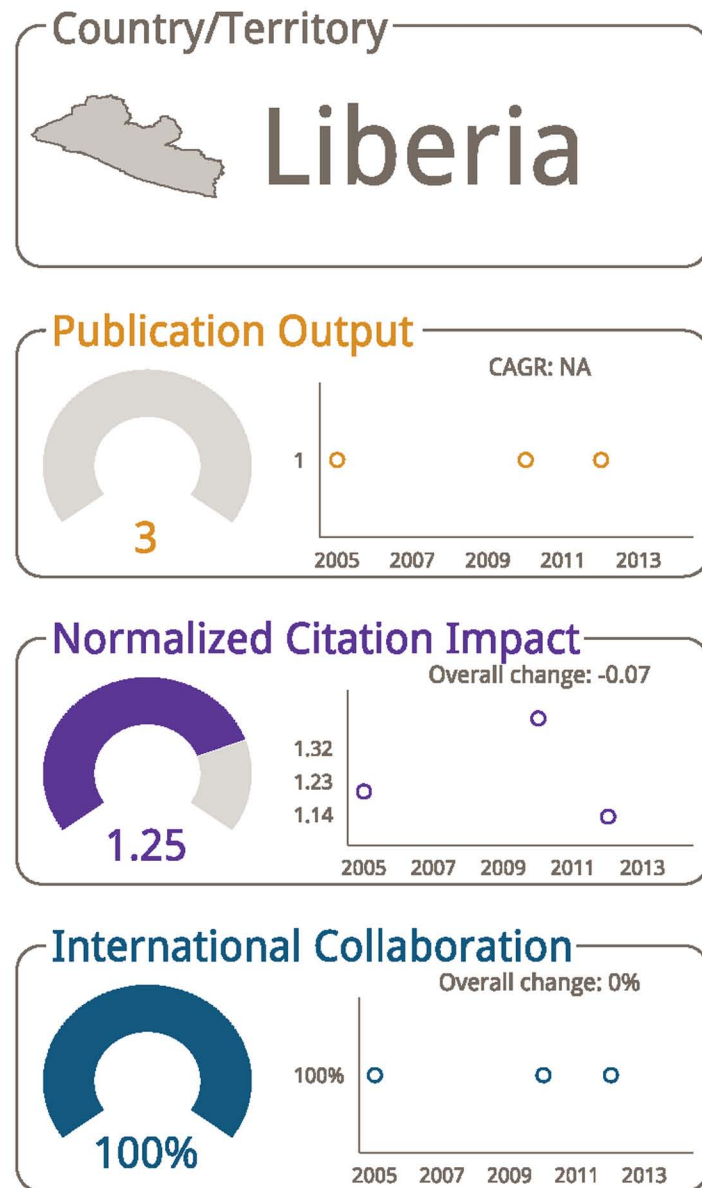


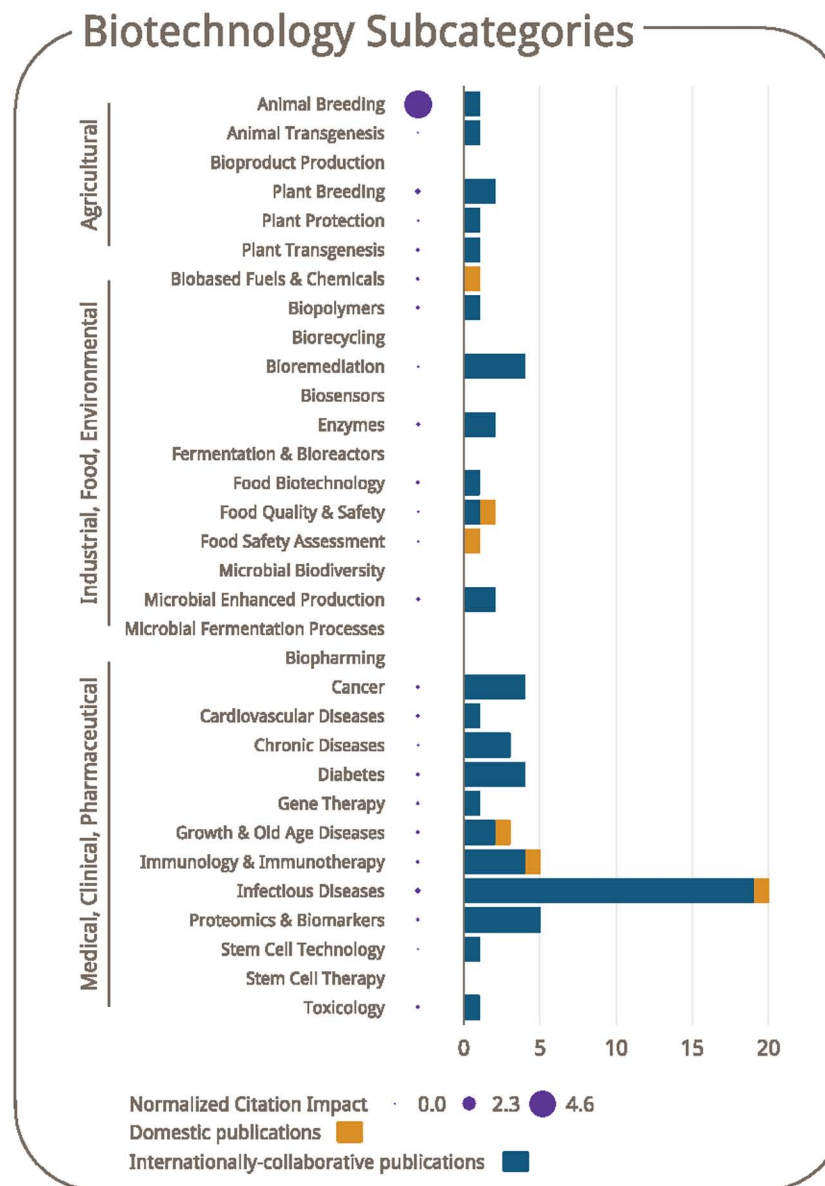
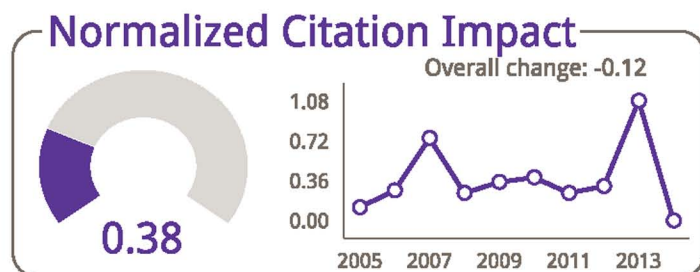
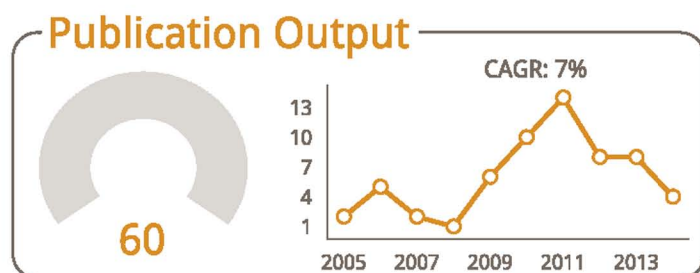




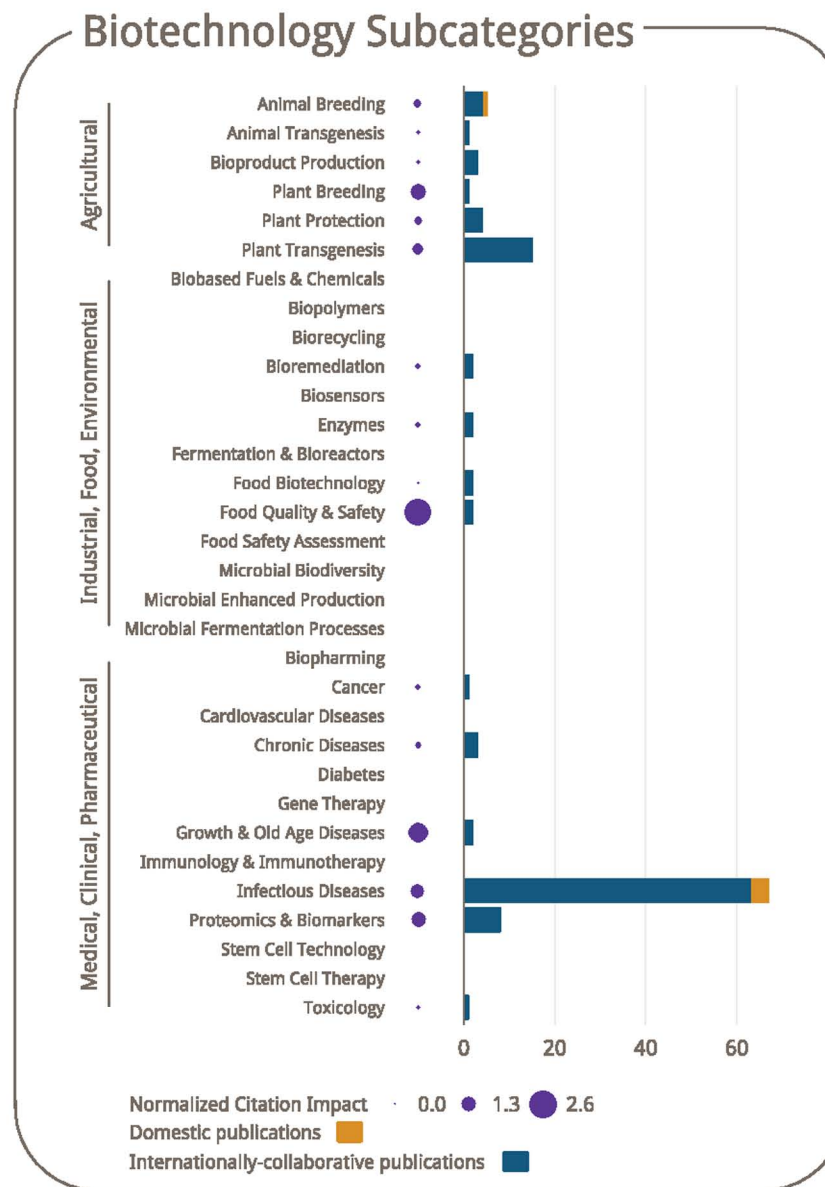
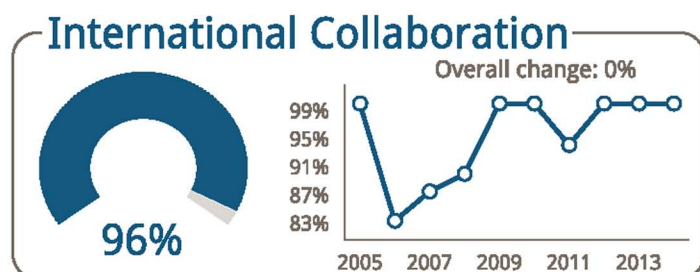
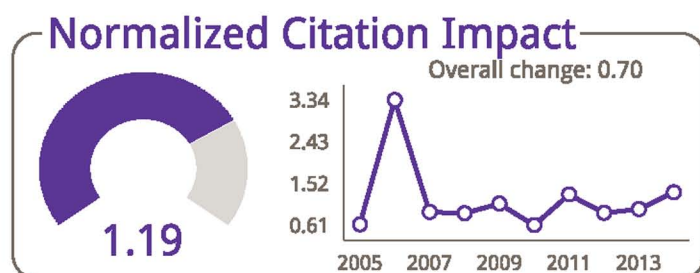
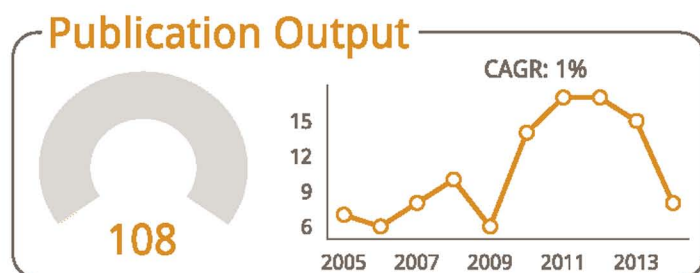


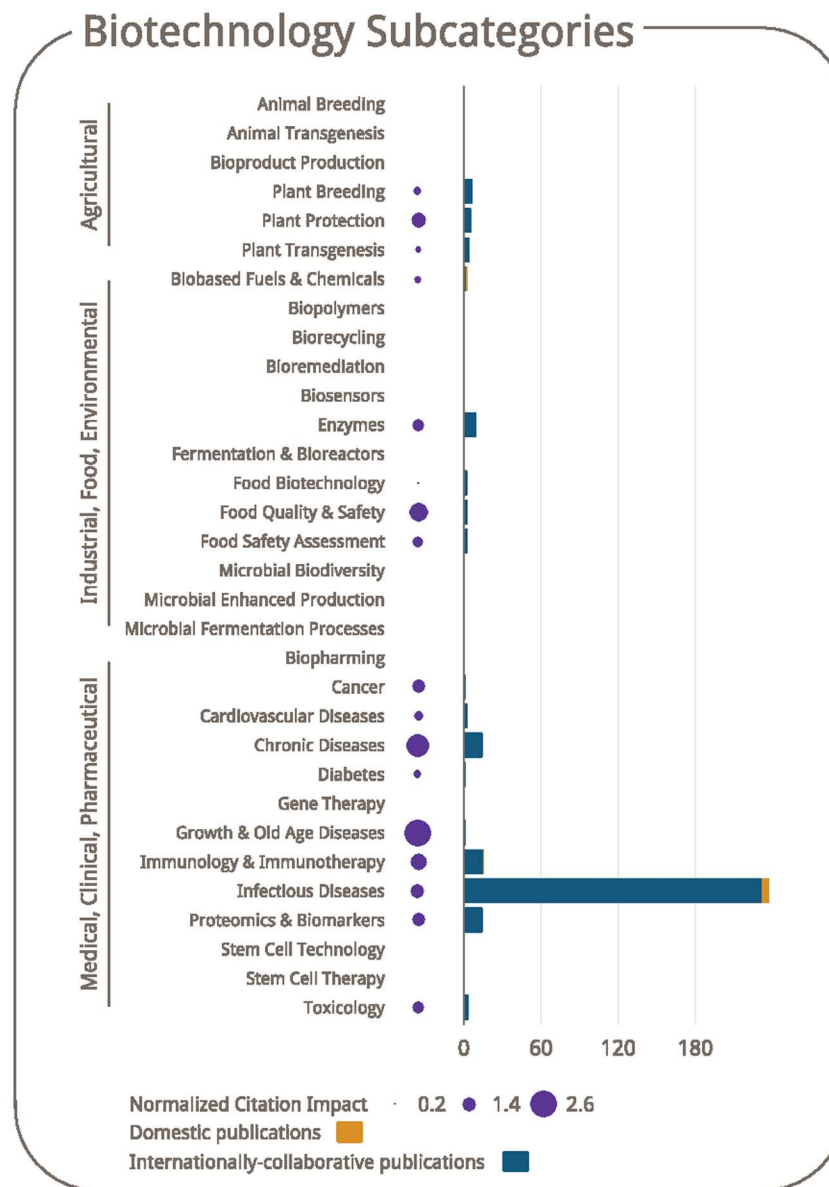
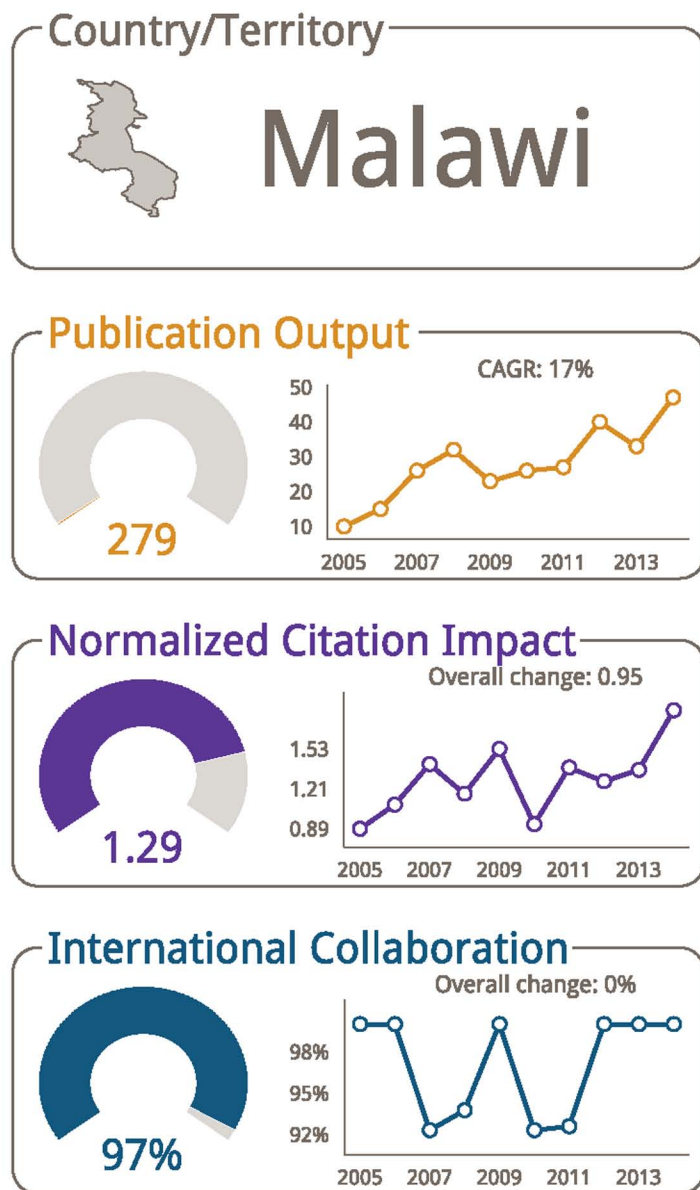


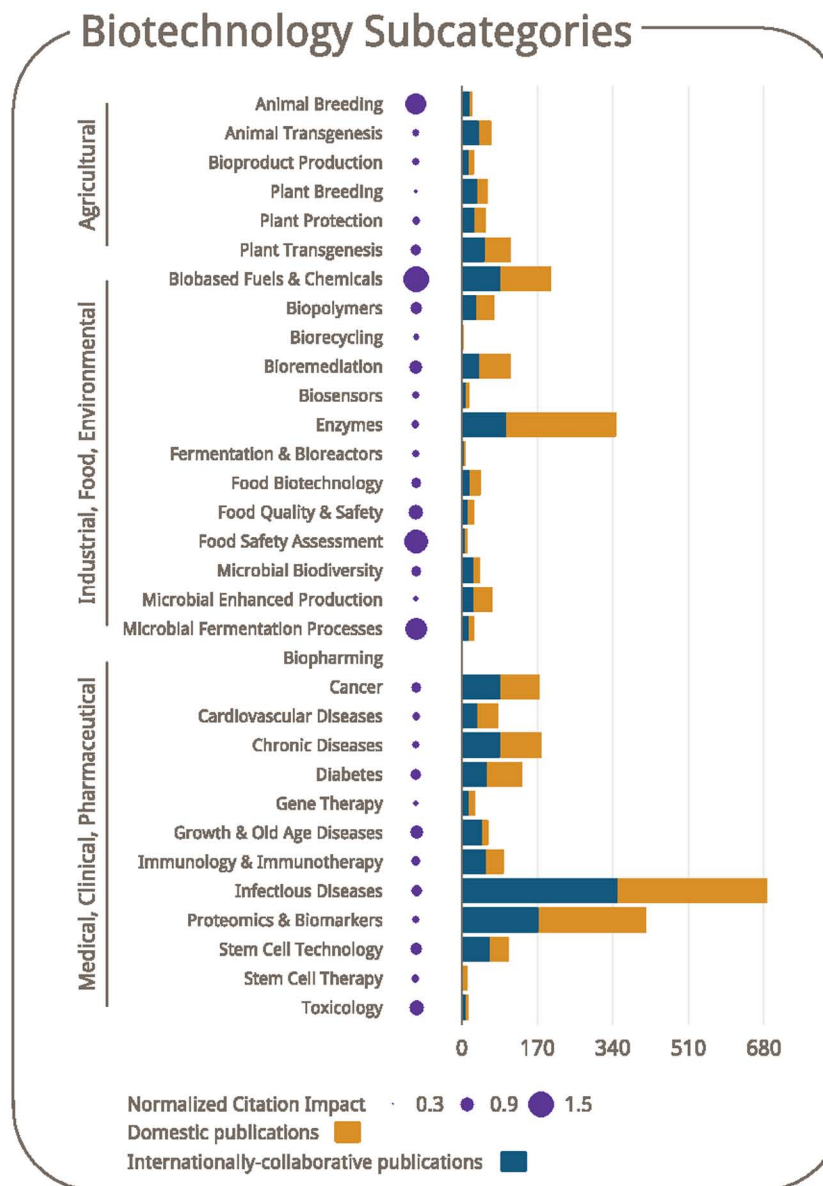
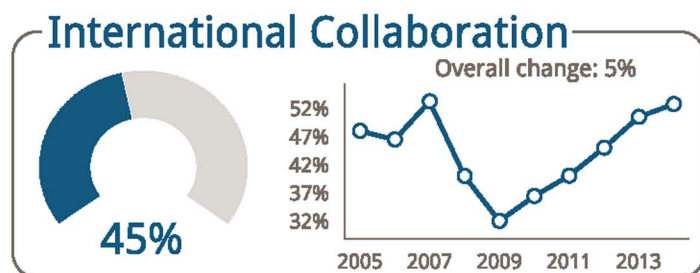
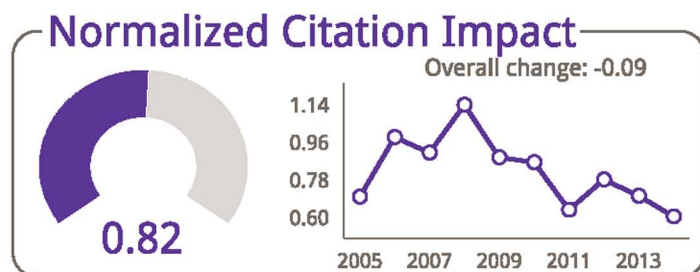
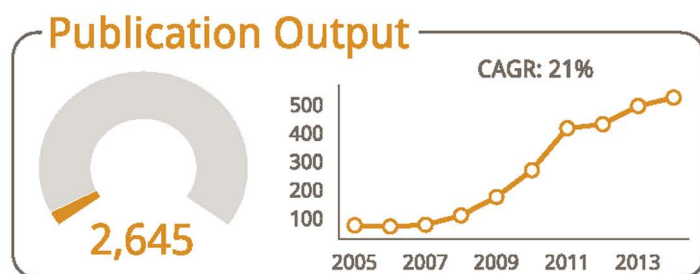


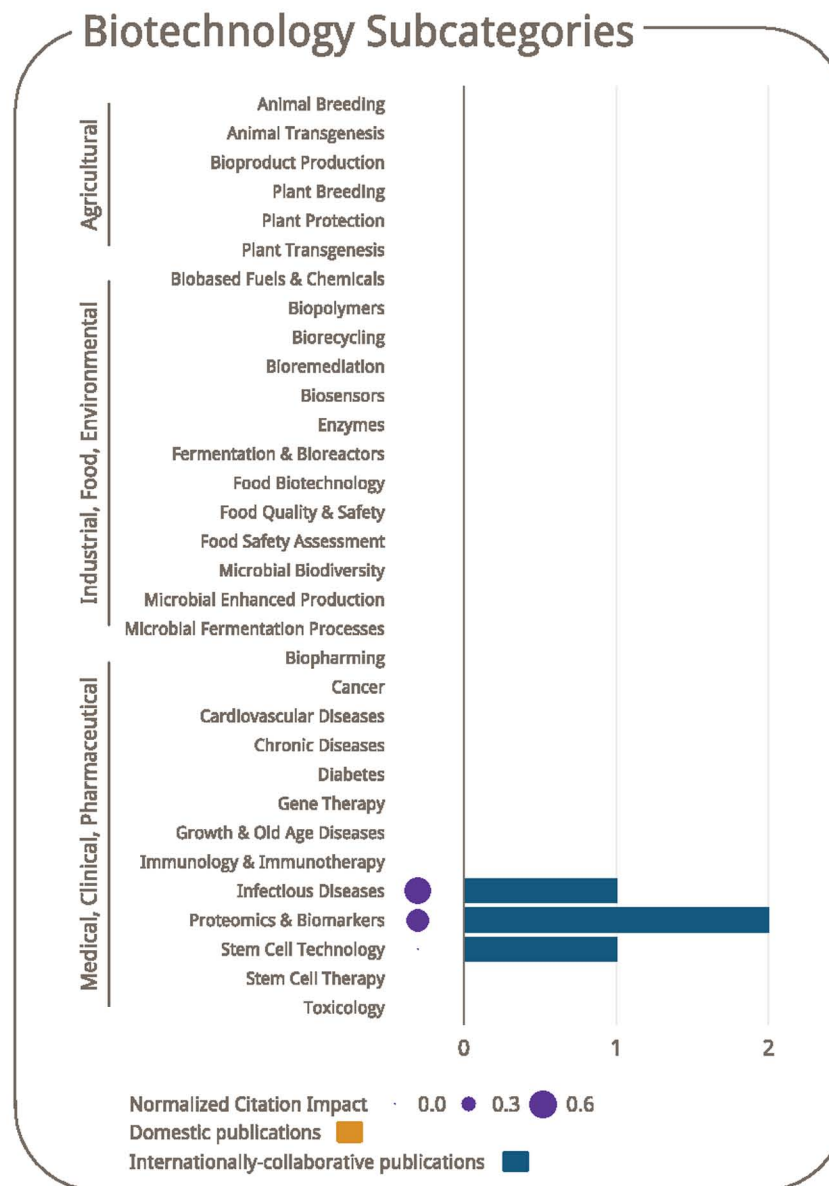
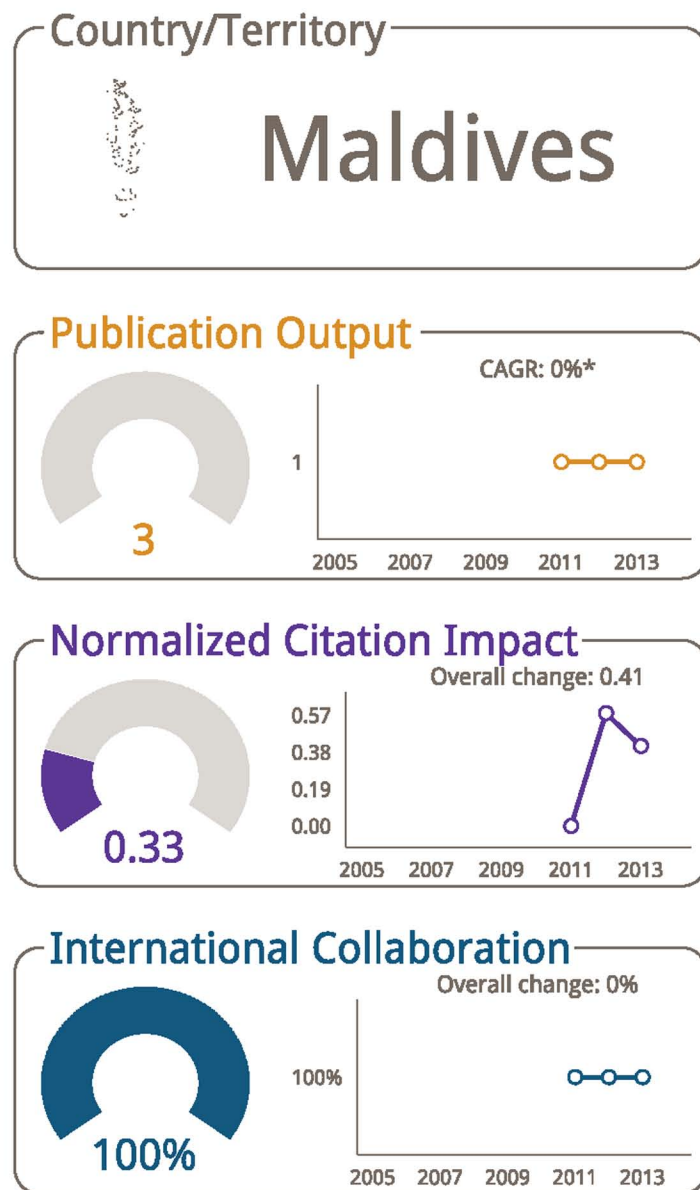


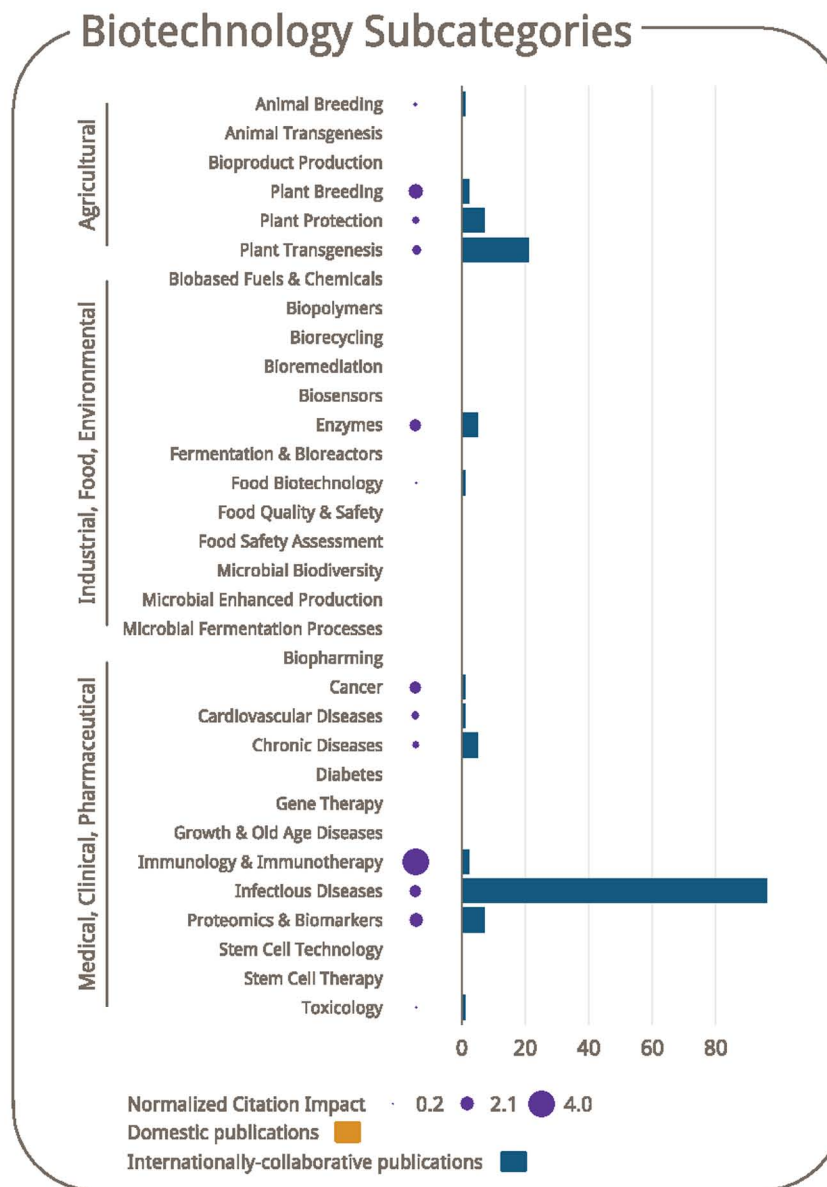
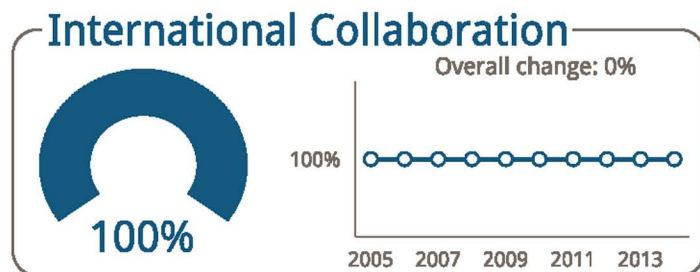
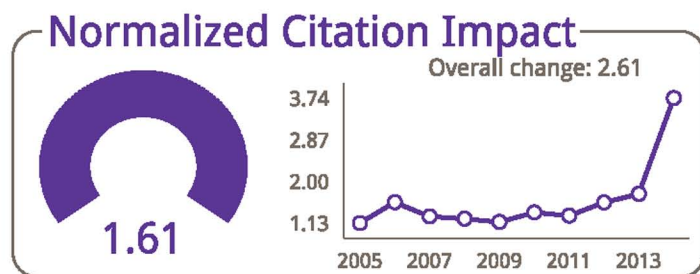
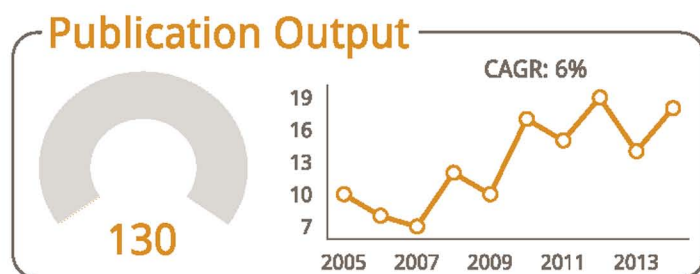




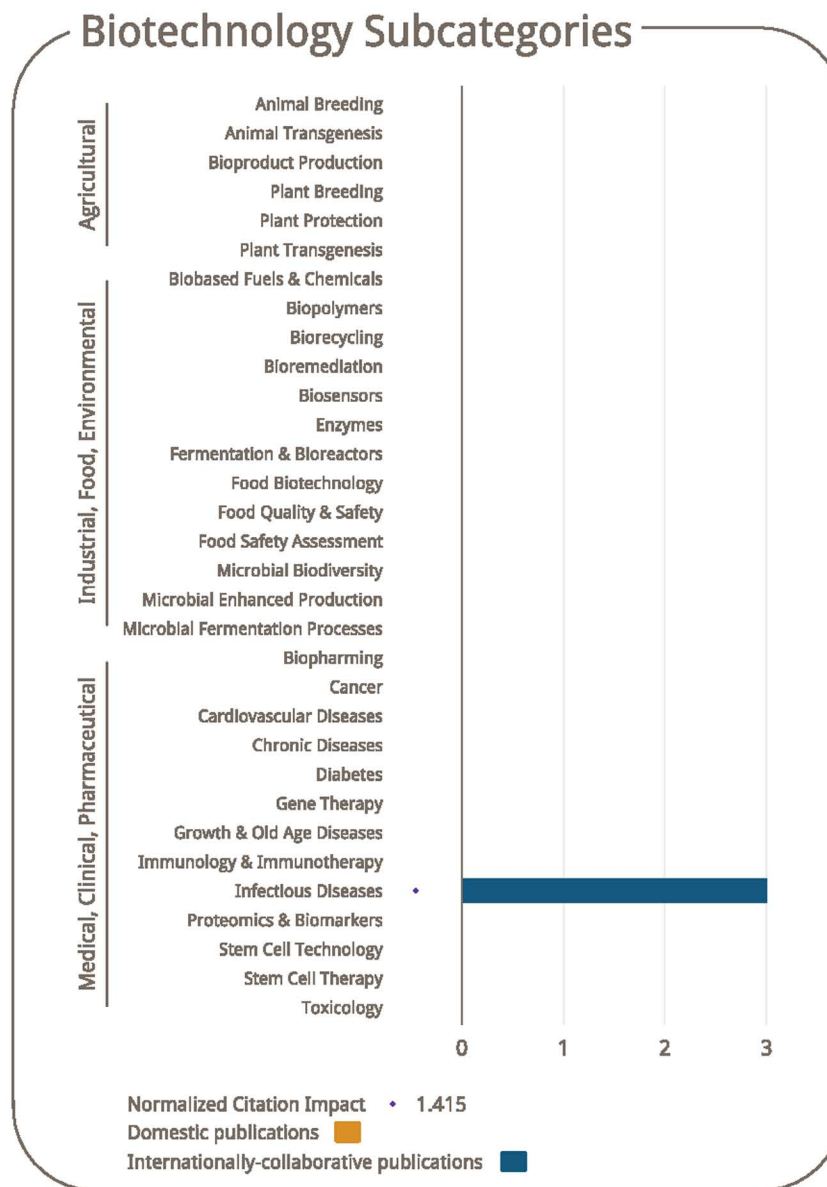
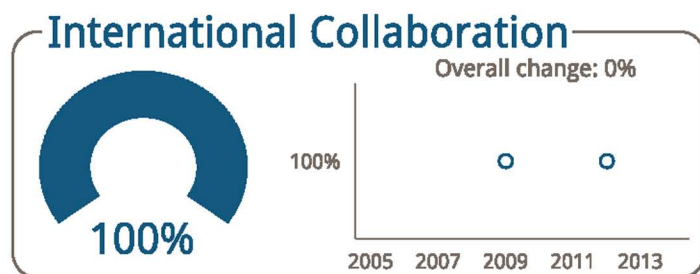
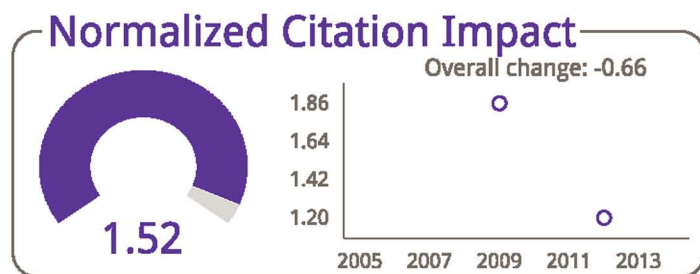
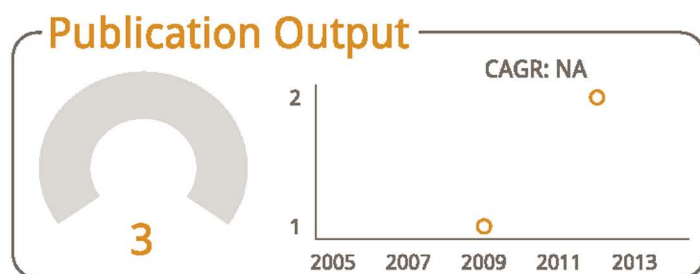










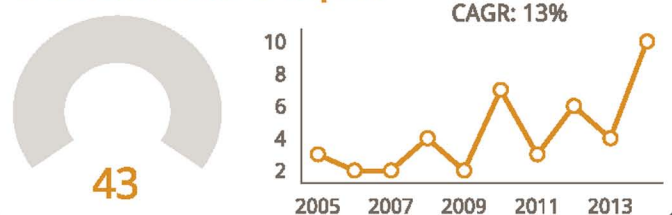




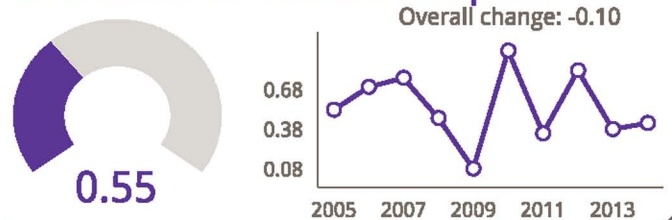
## Country/Territory

# Mauritius

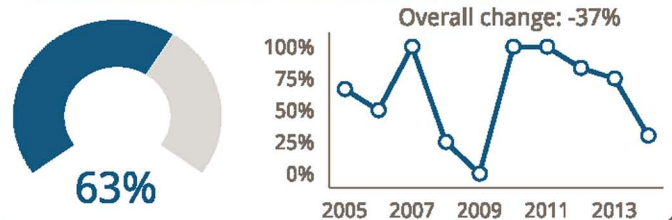
## Publication Output



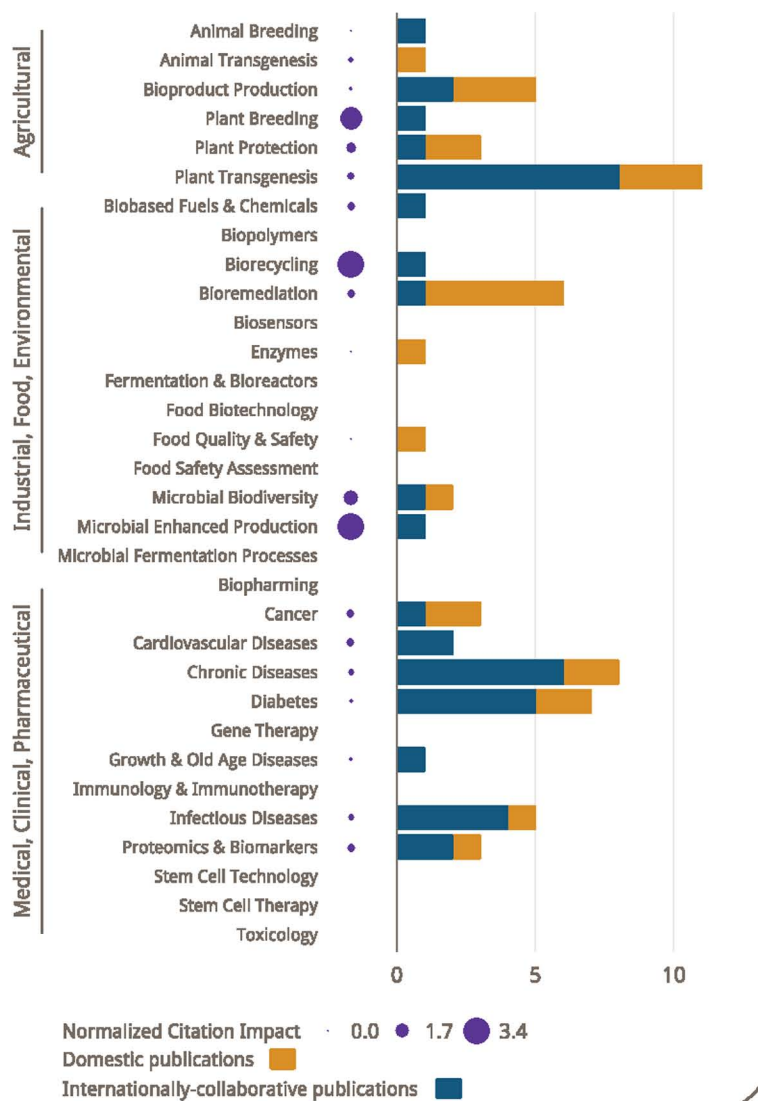
## Normalized Citation Impact

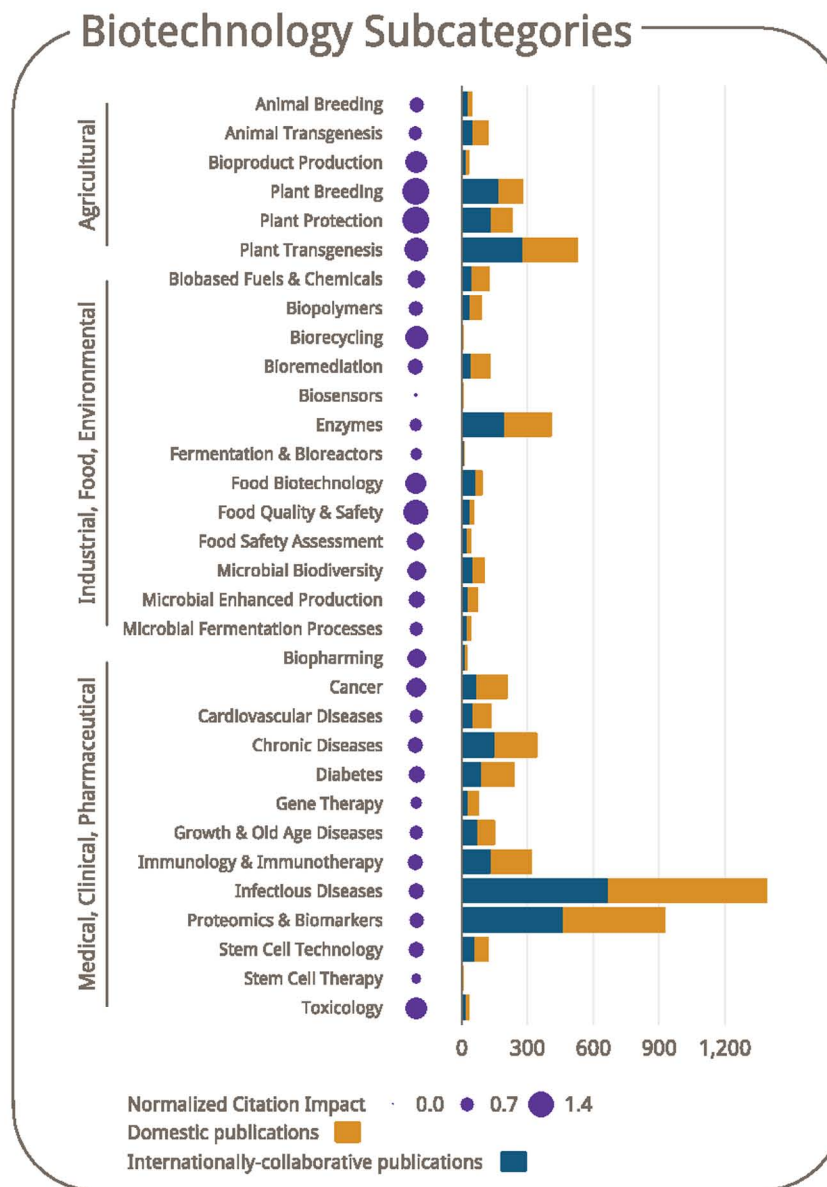
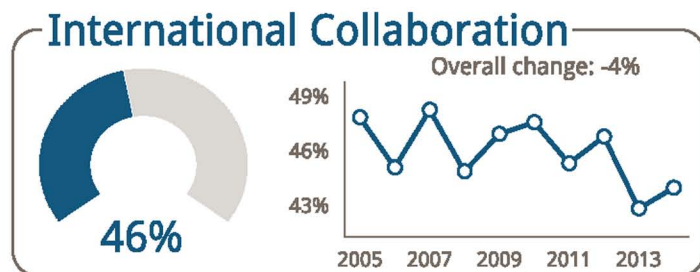
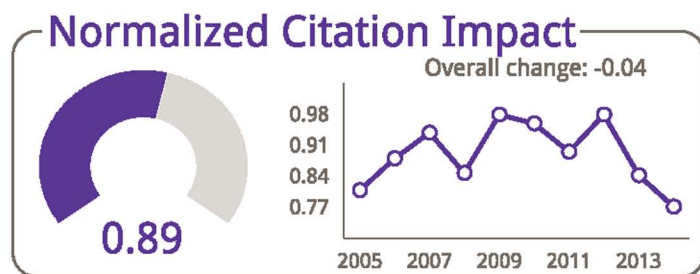
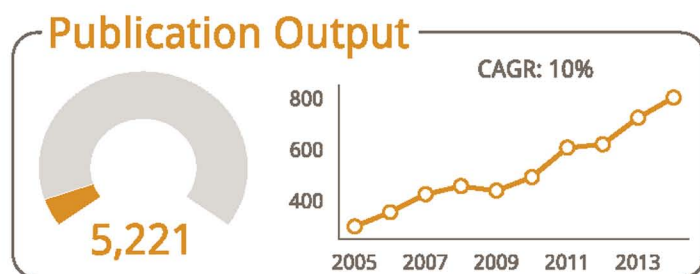


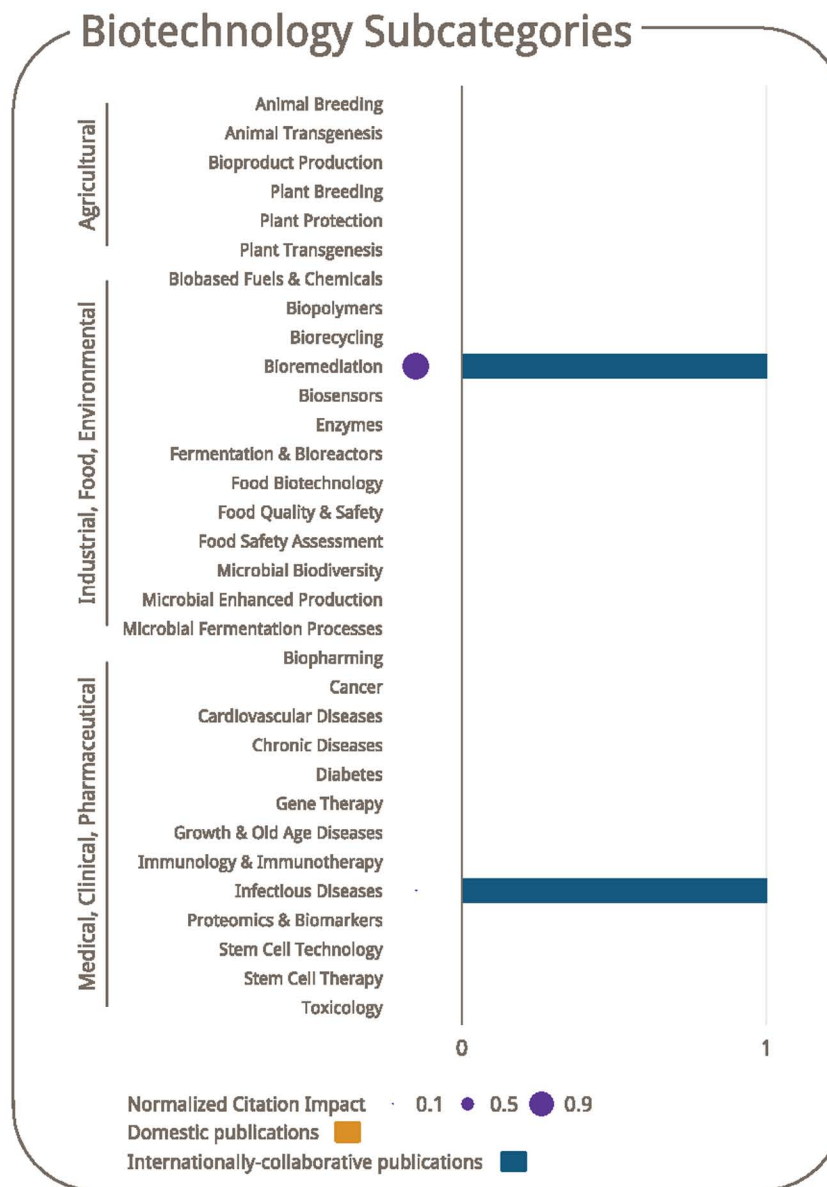
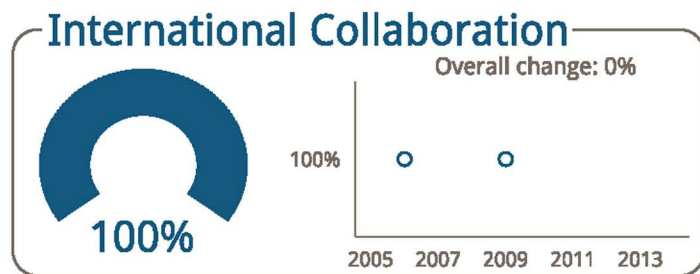
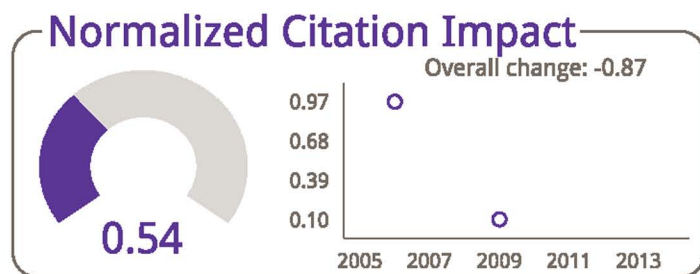
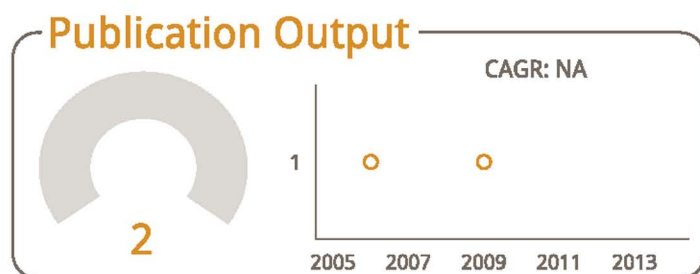
## International Collaboration

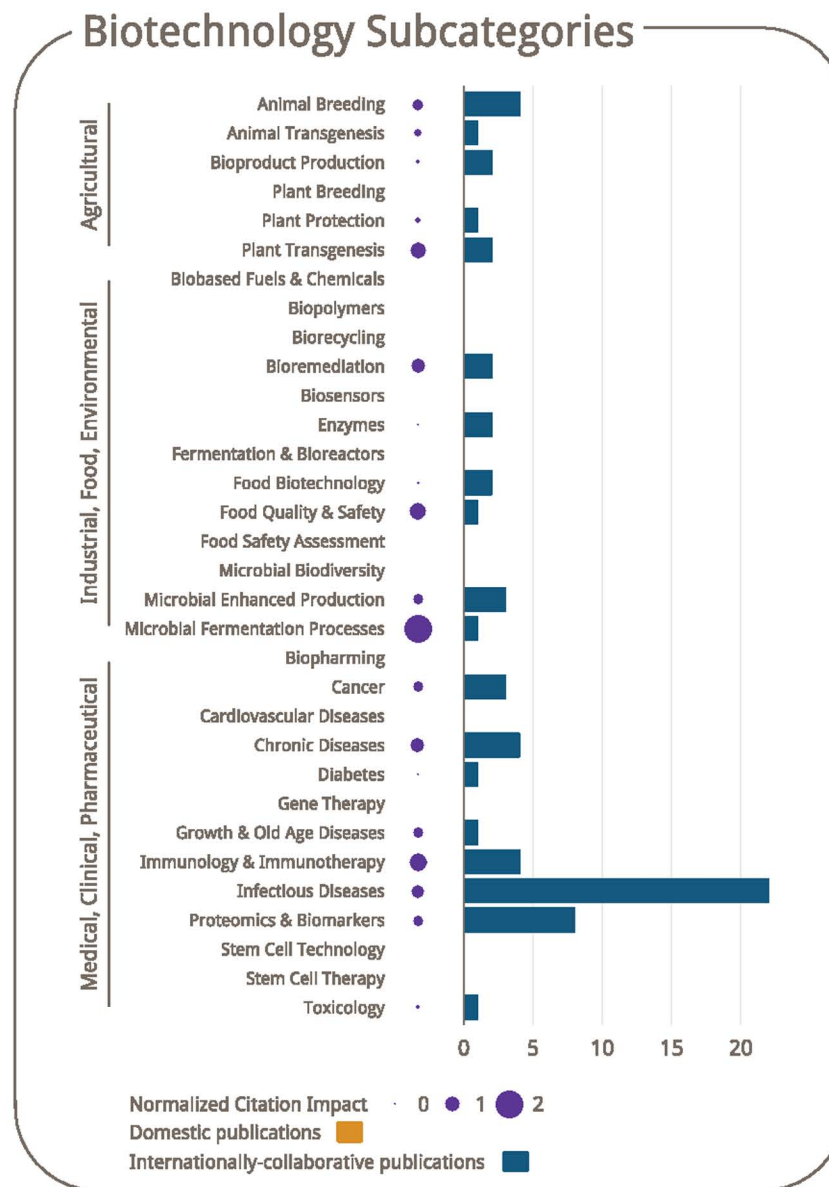
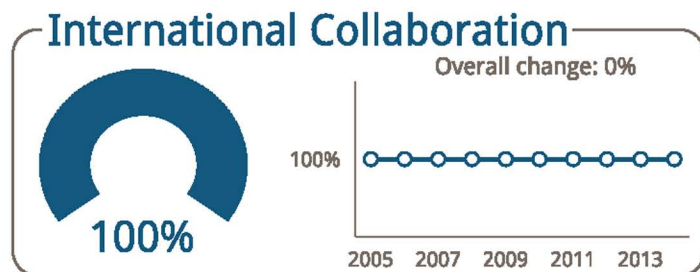
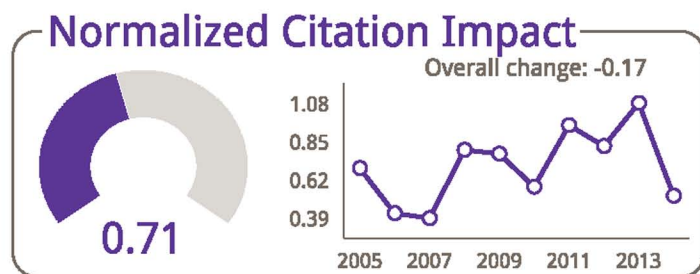
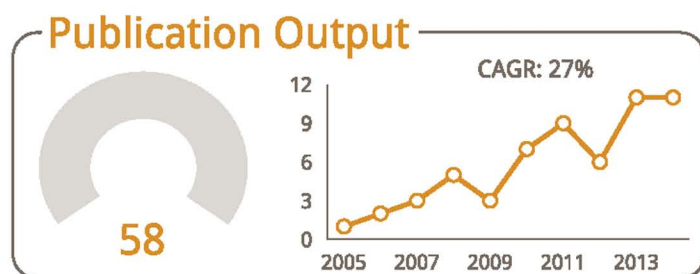


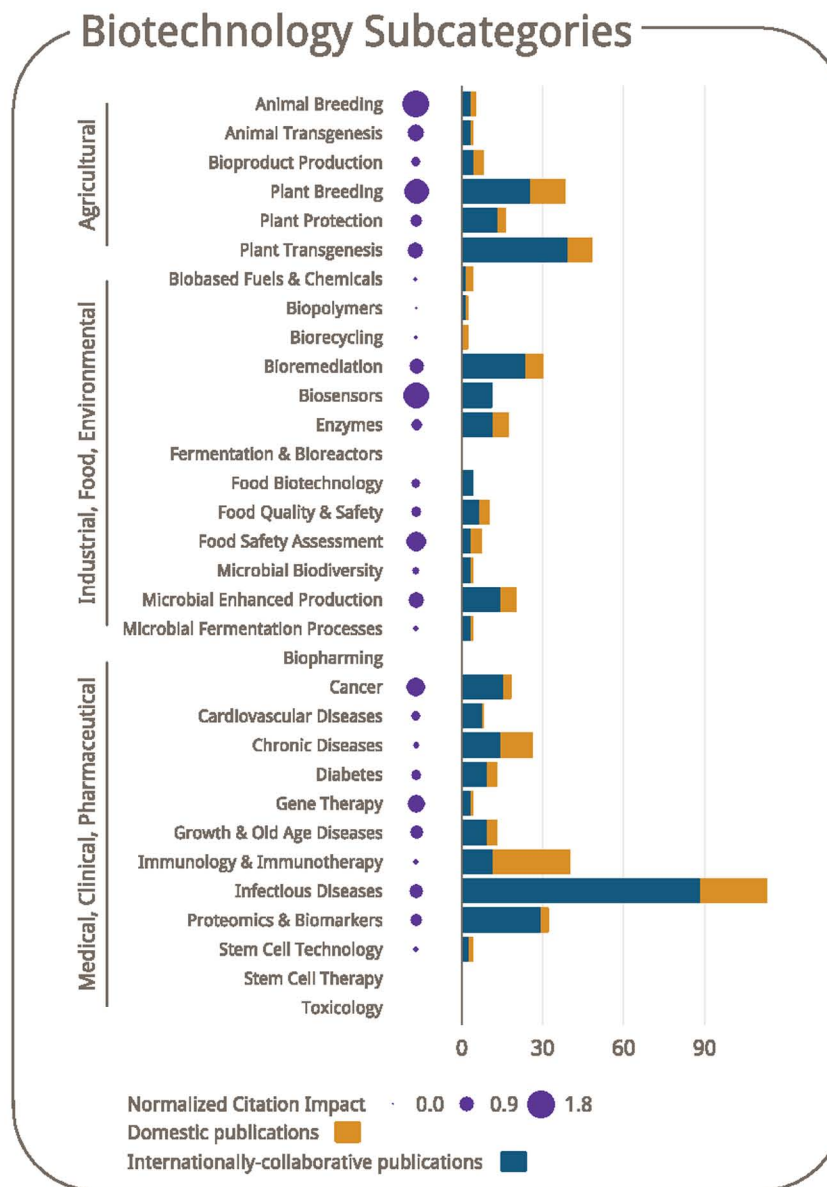
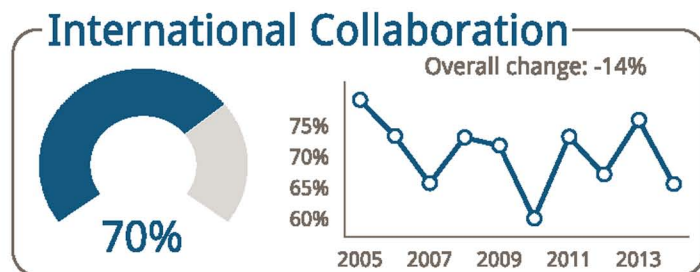
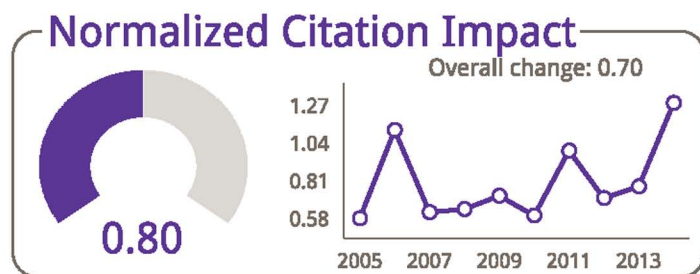
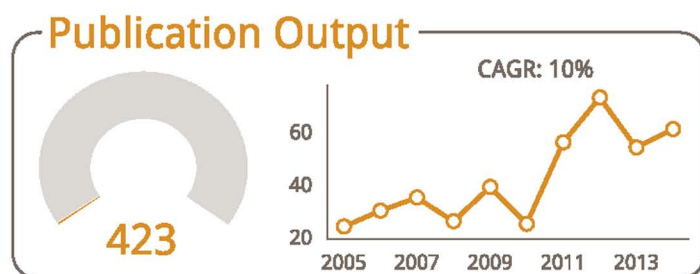
## Biotechnology Subcategories



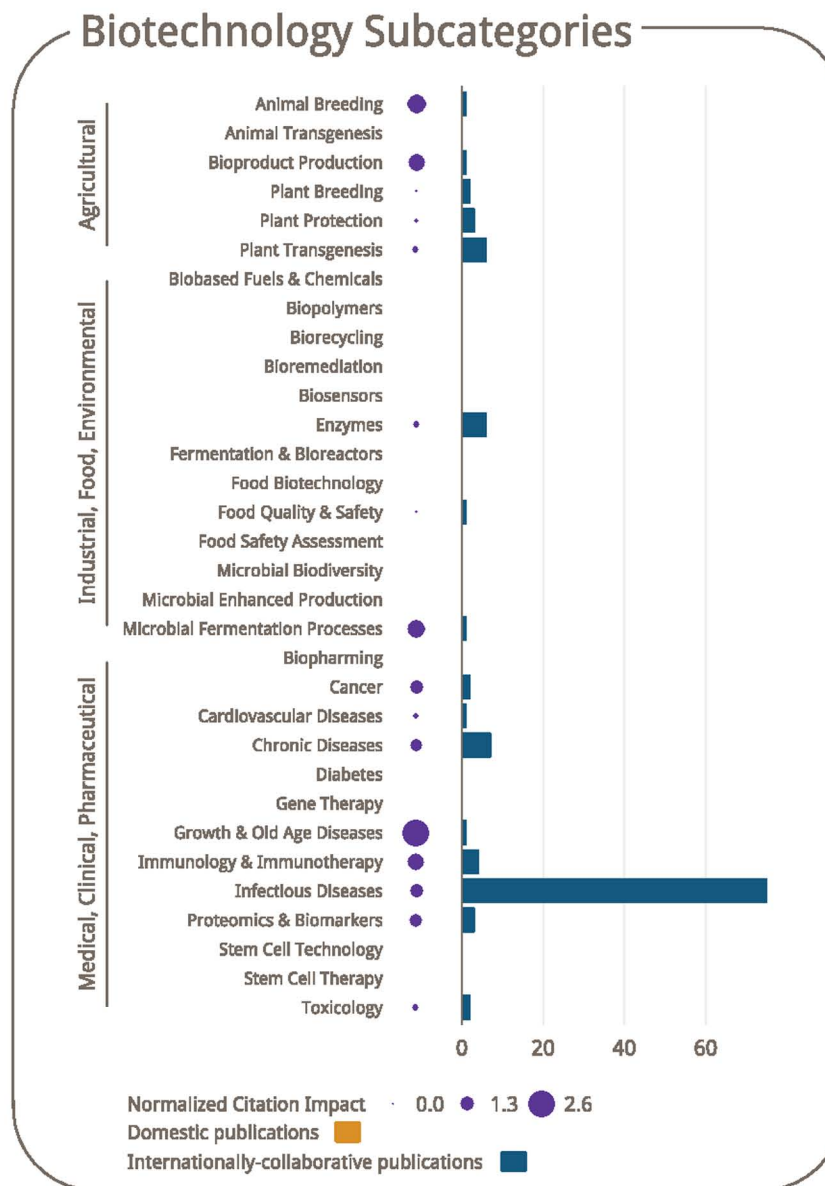
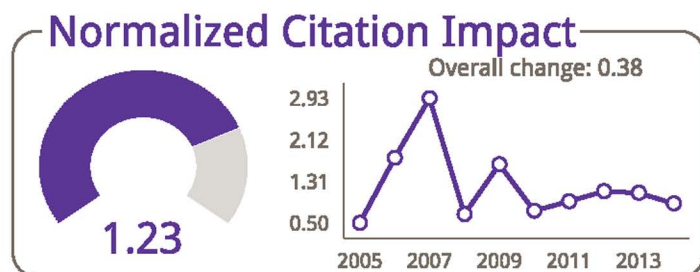
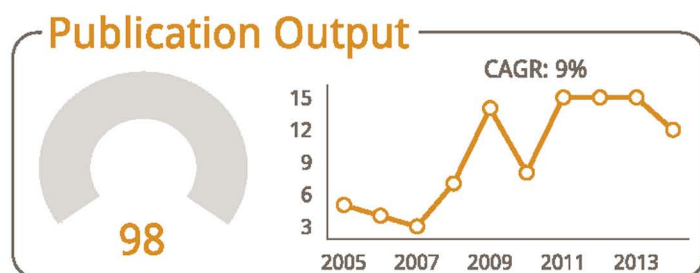




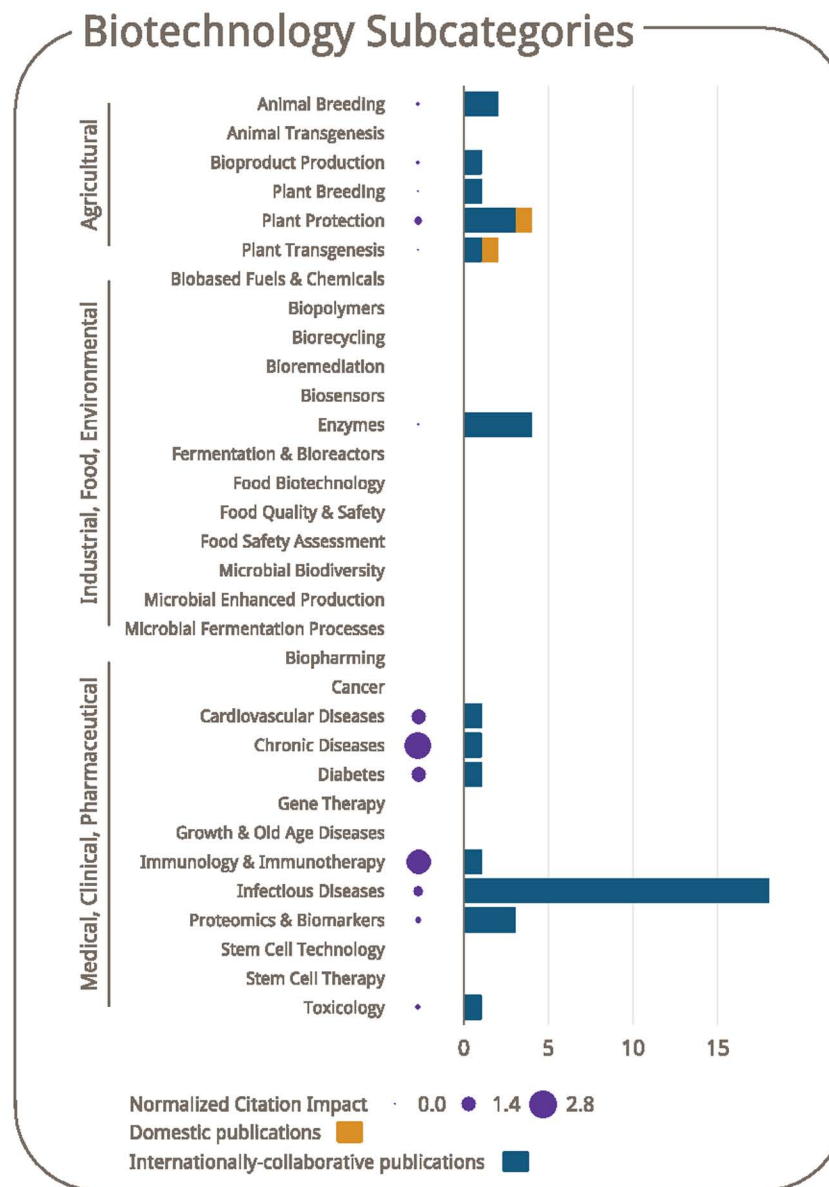
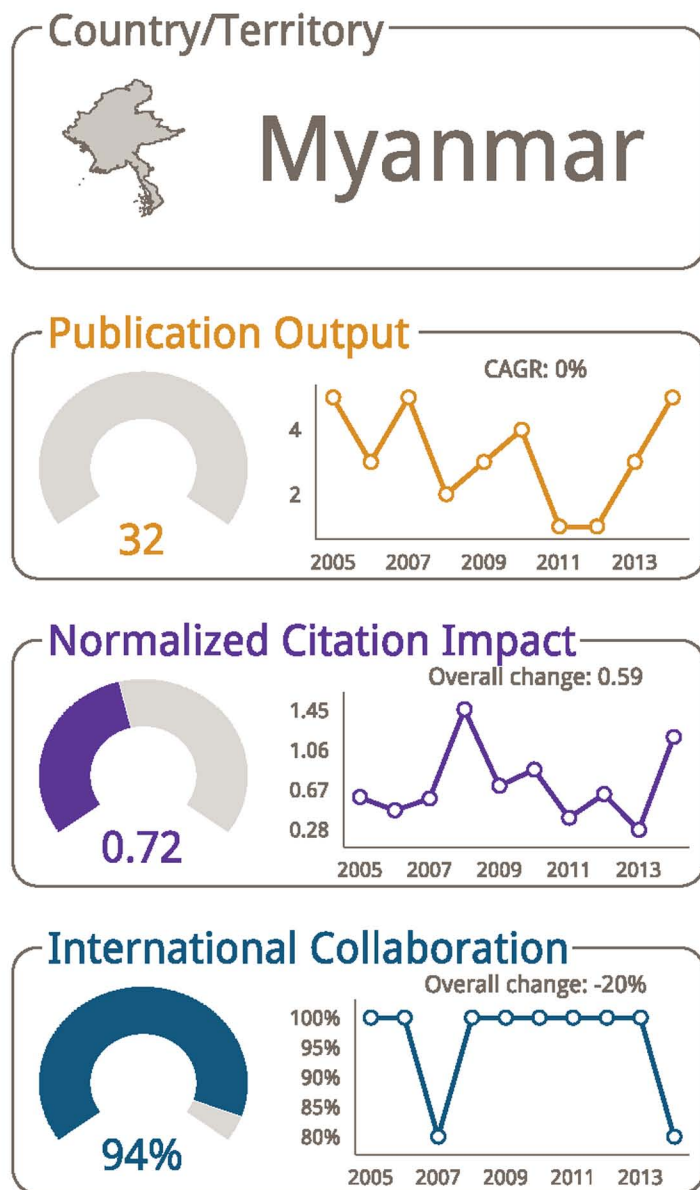


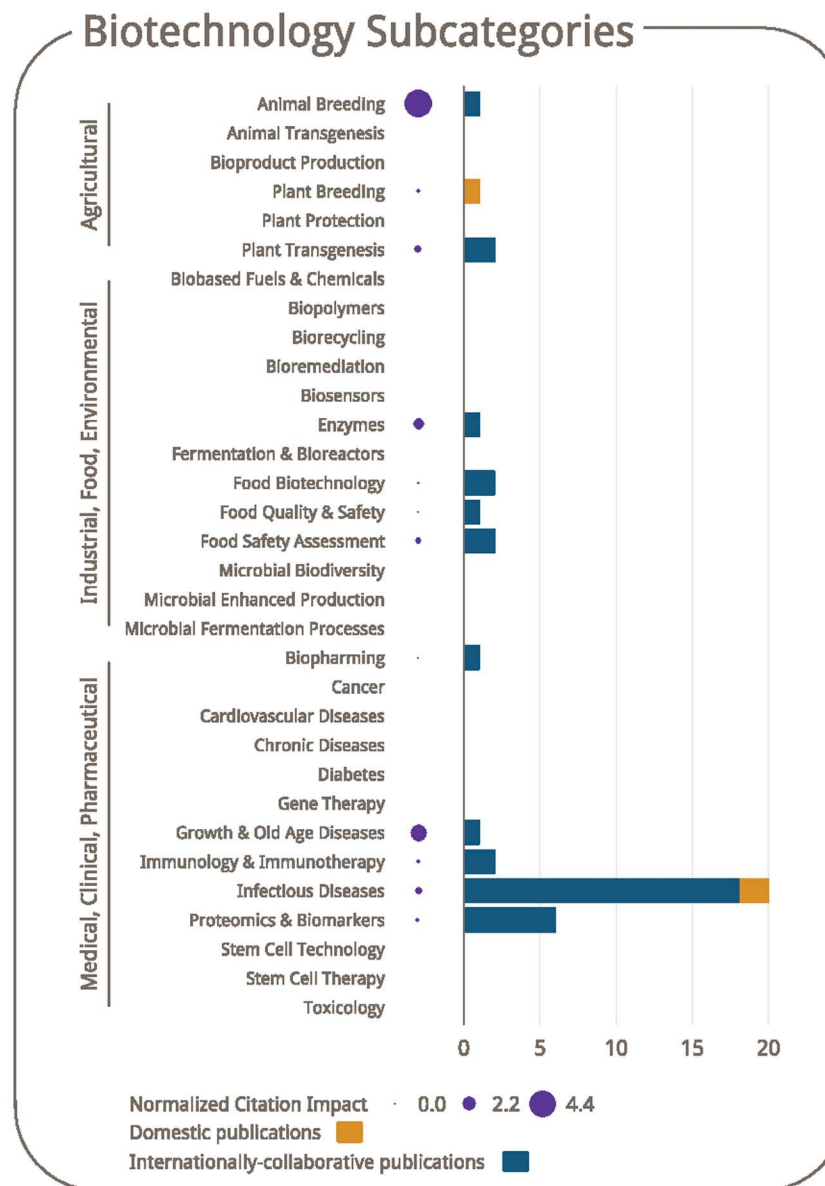
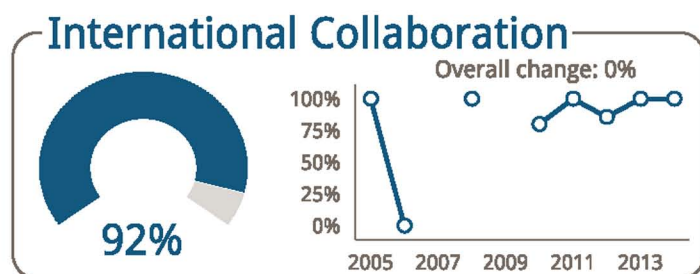
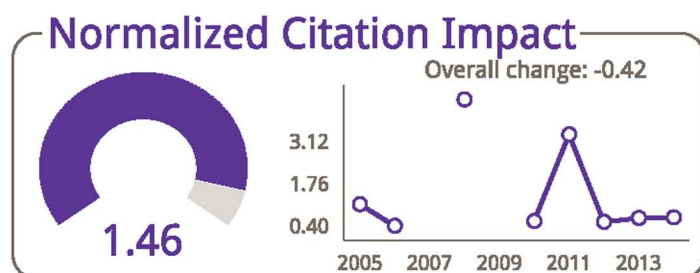
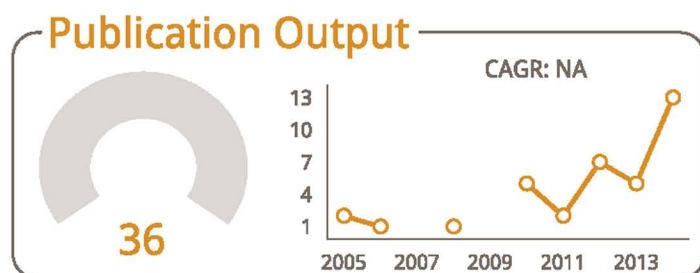


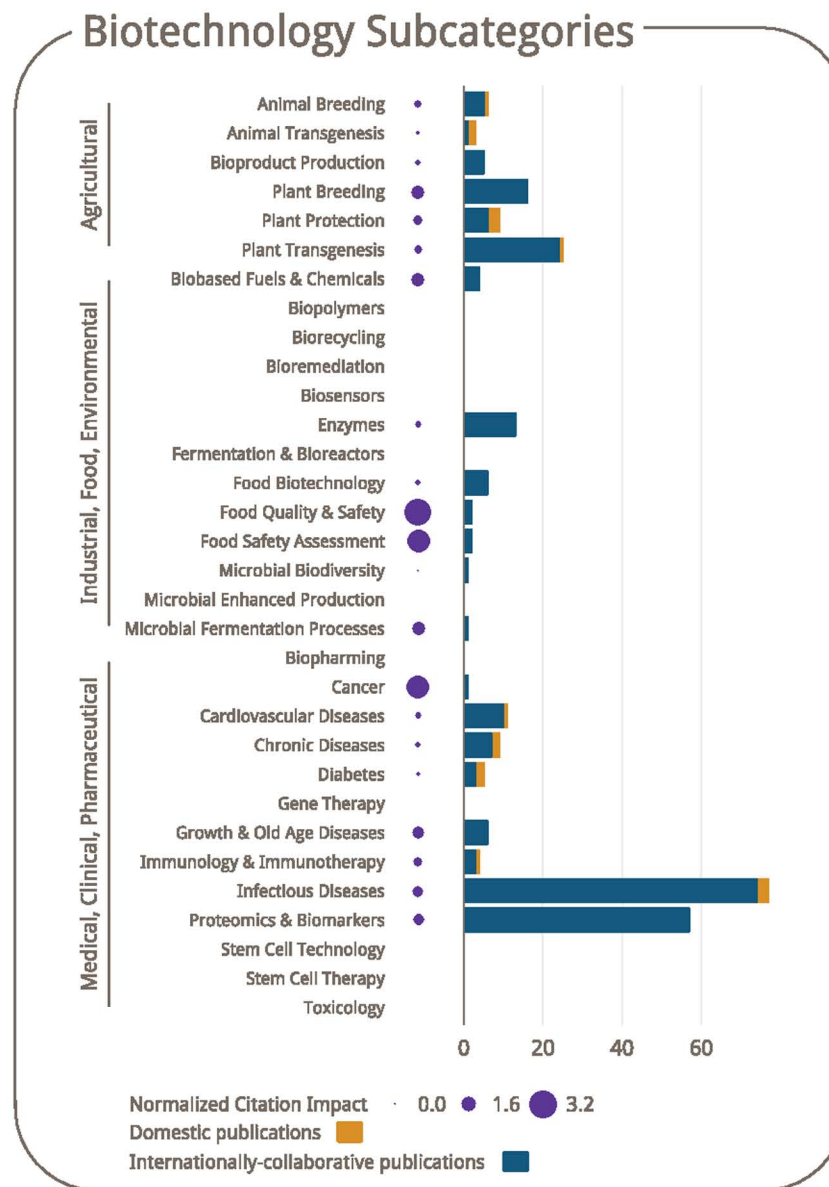
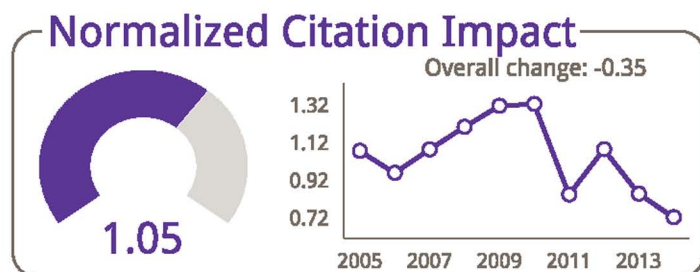
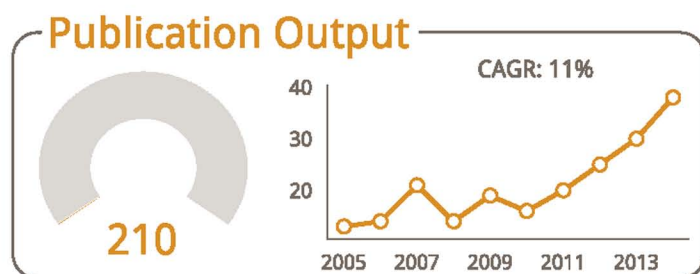


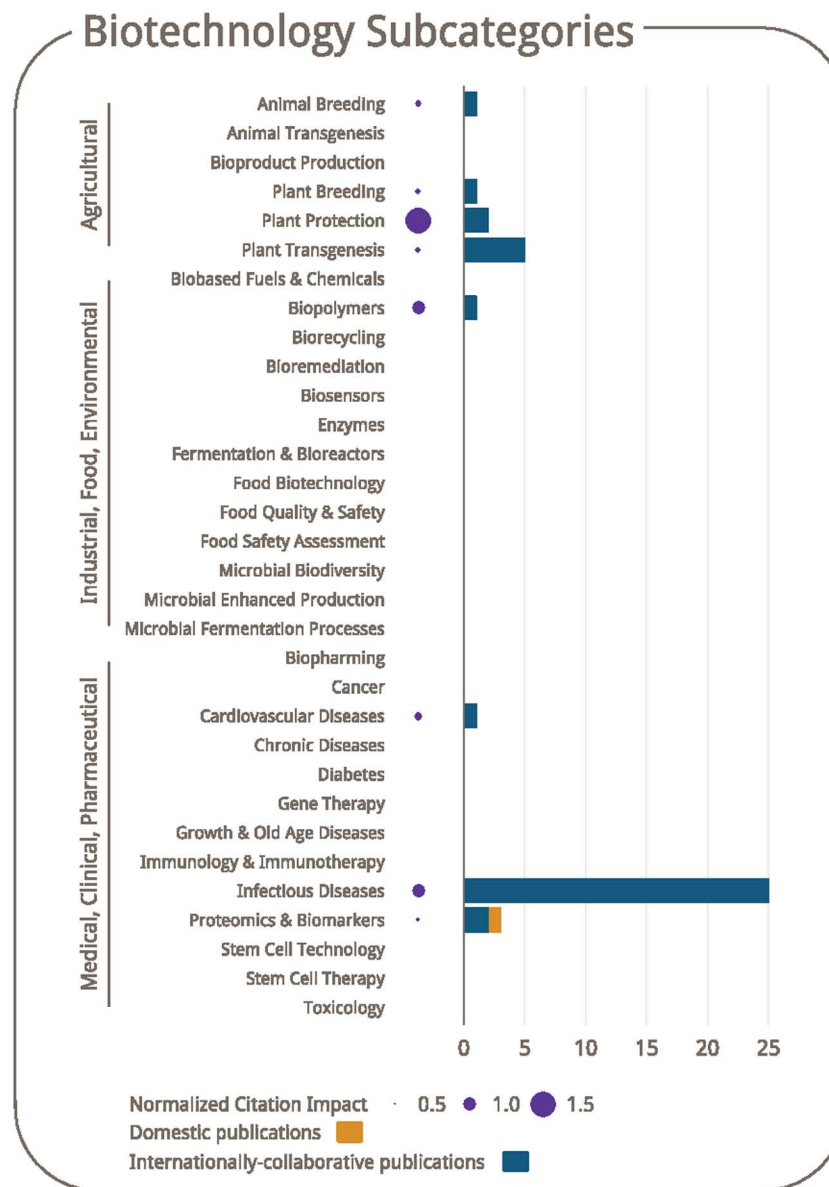
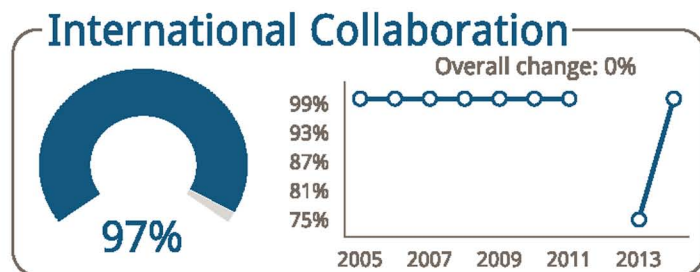
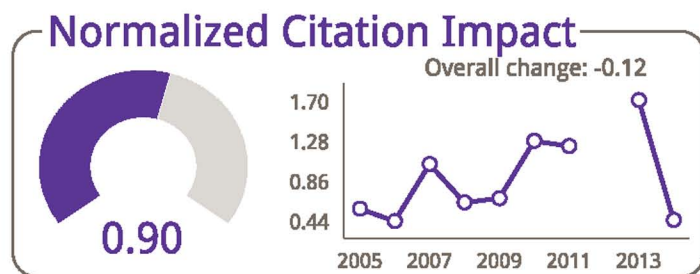
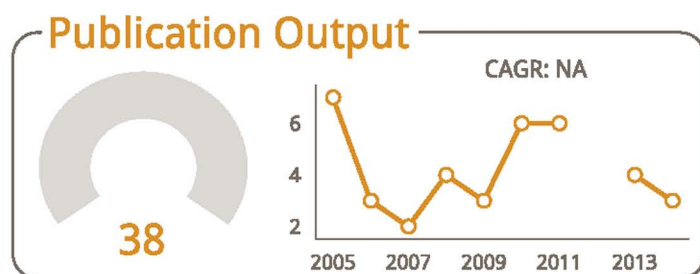


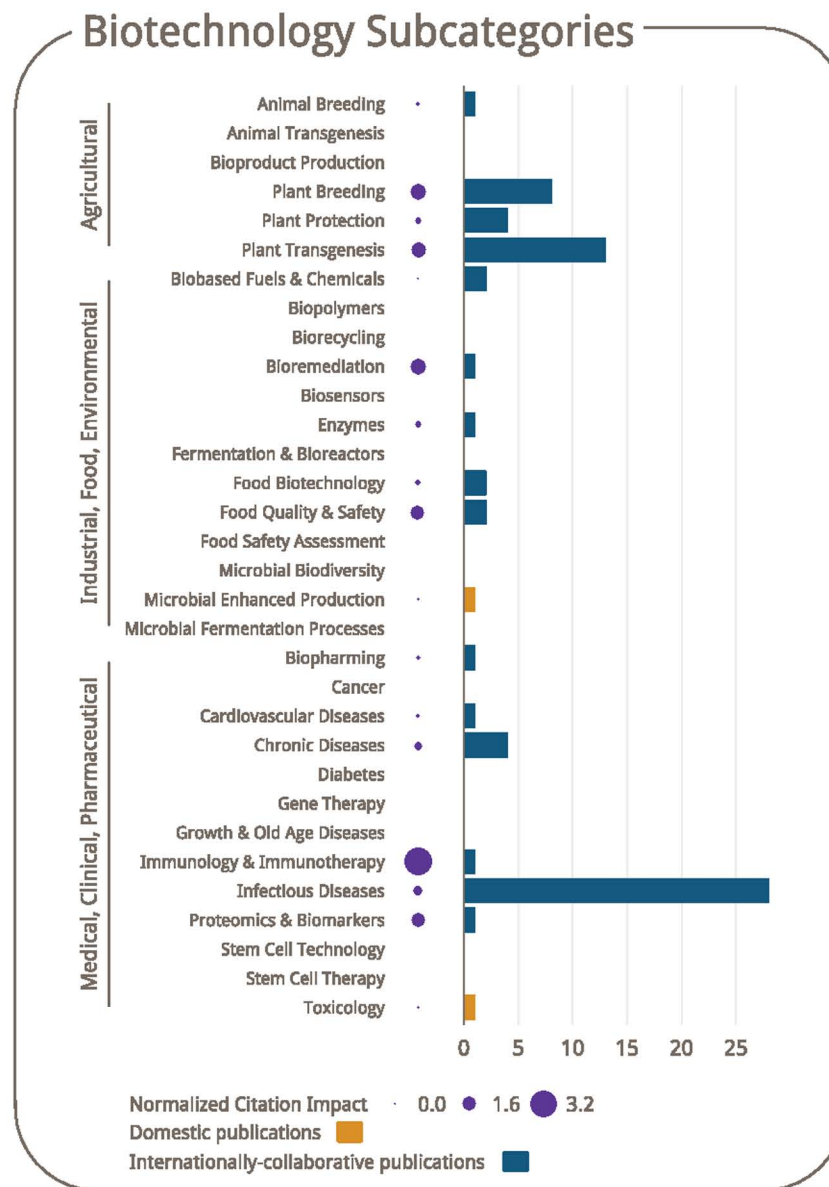
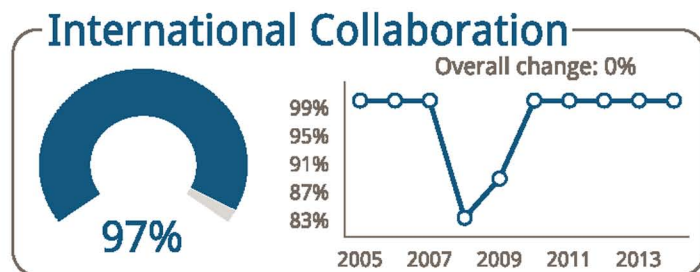
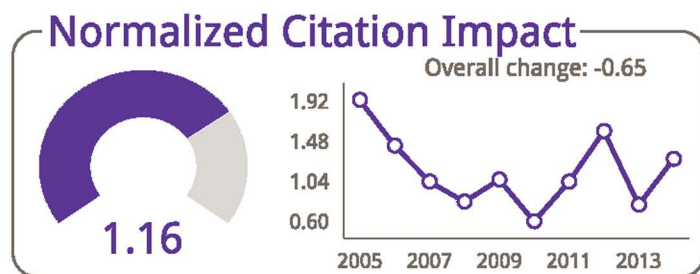
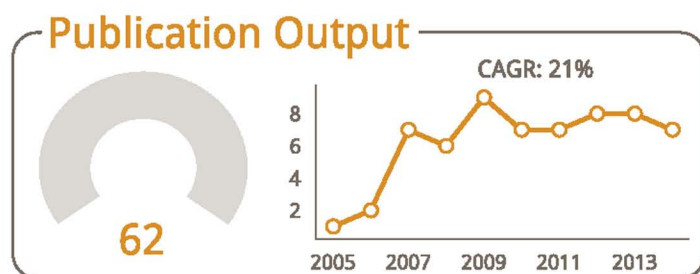




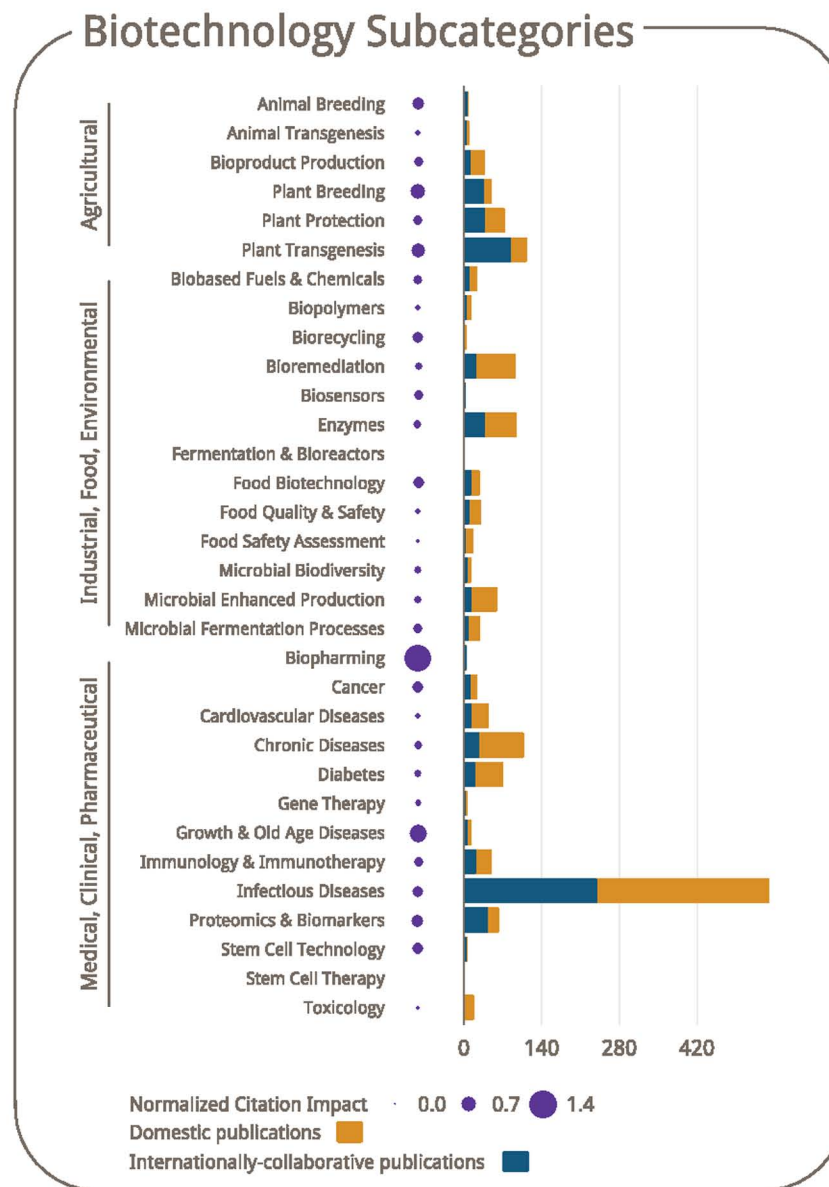
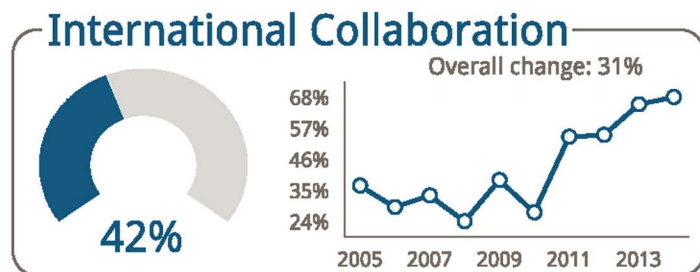
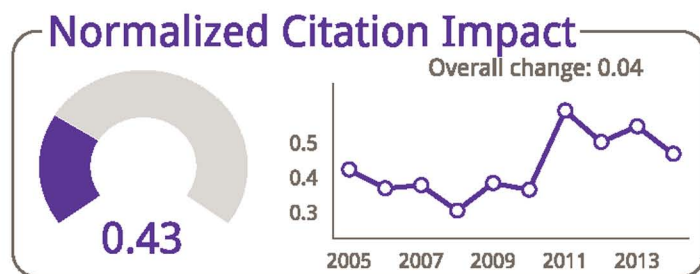
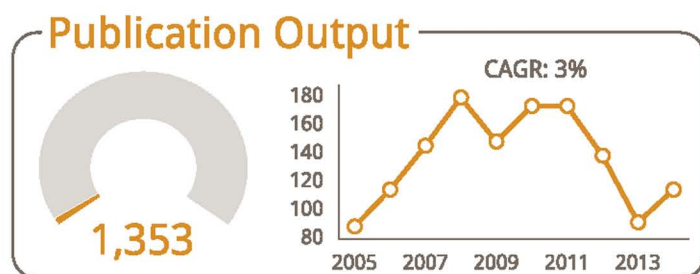




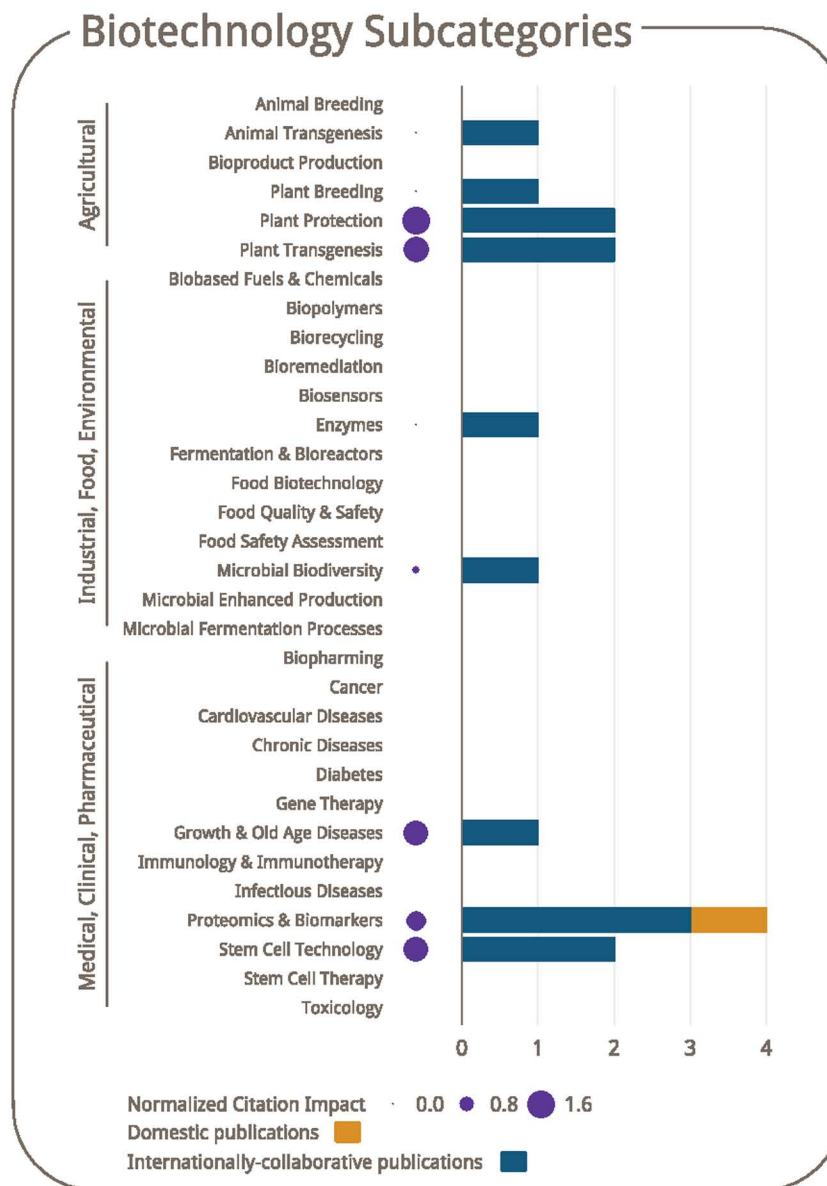
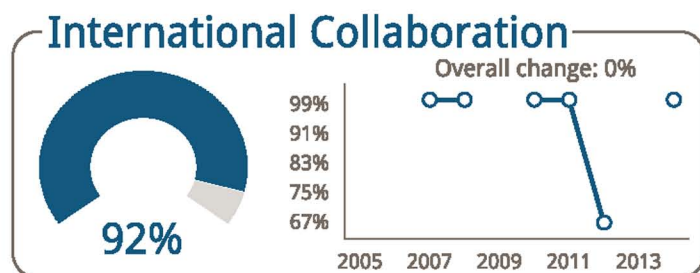
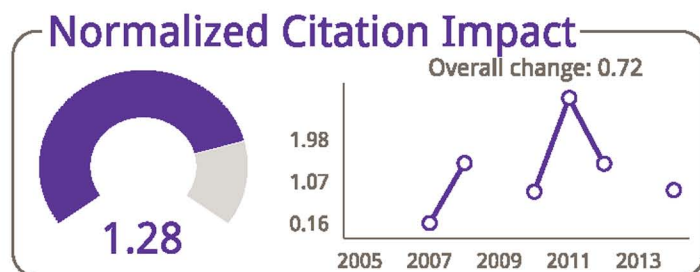
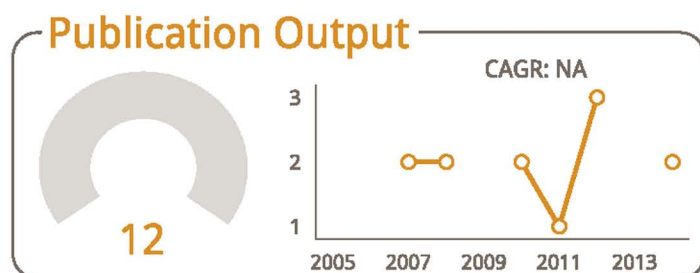


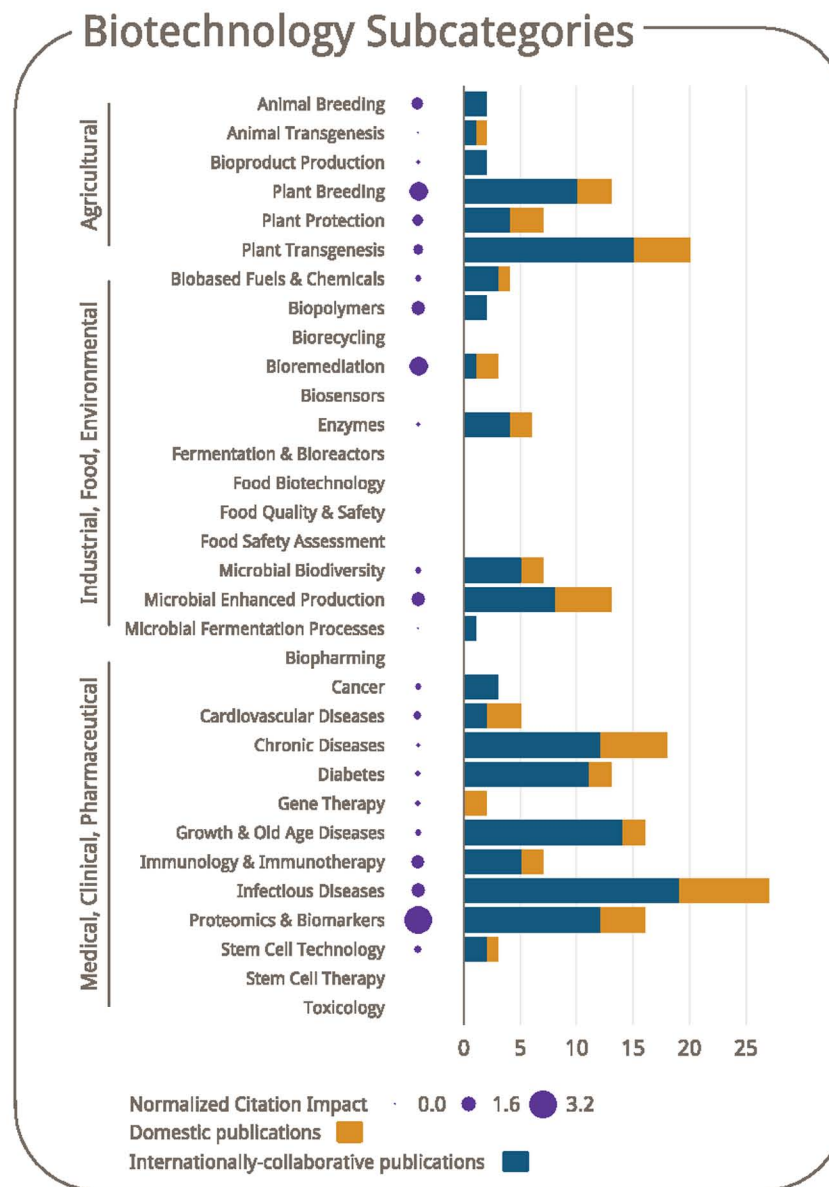
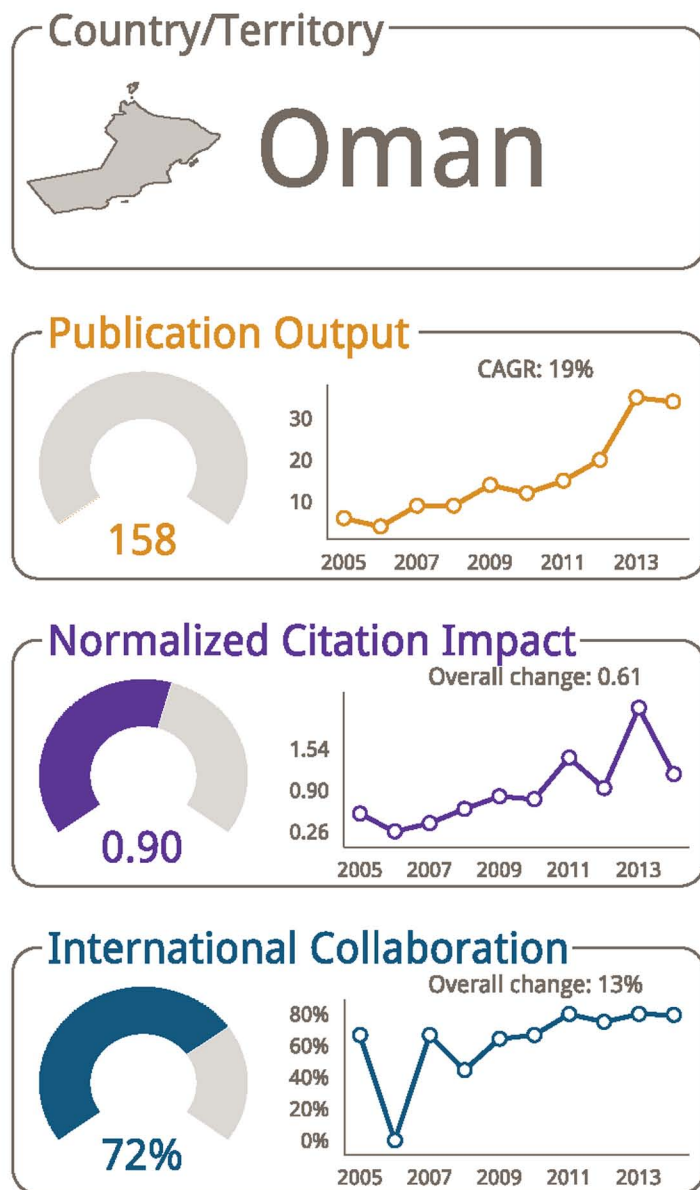


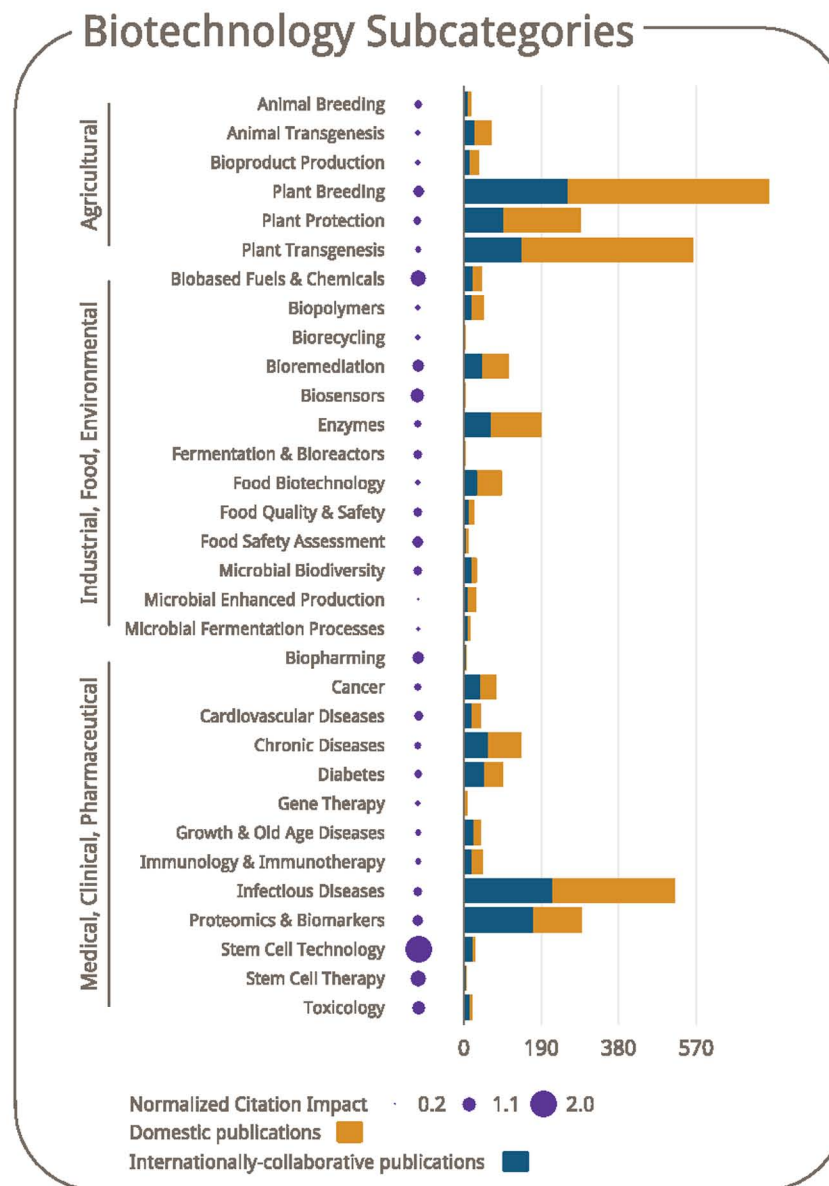
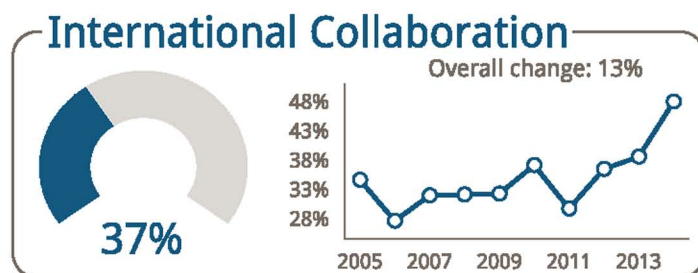
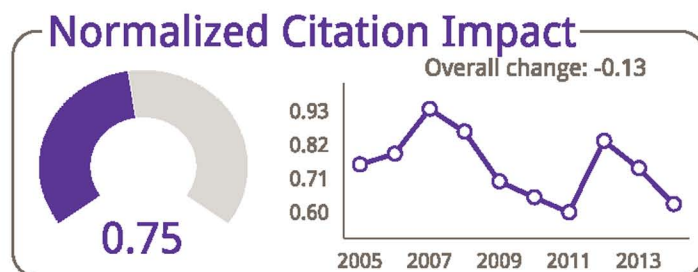
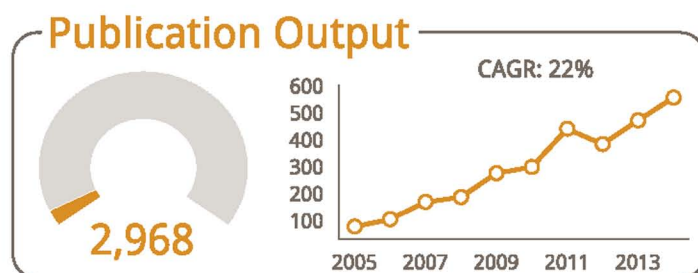


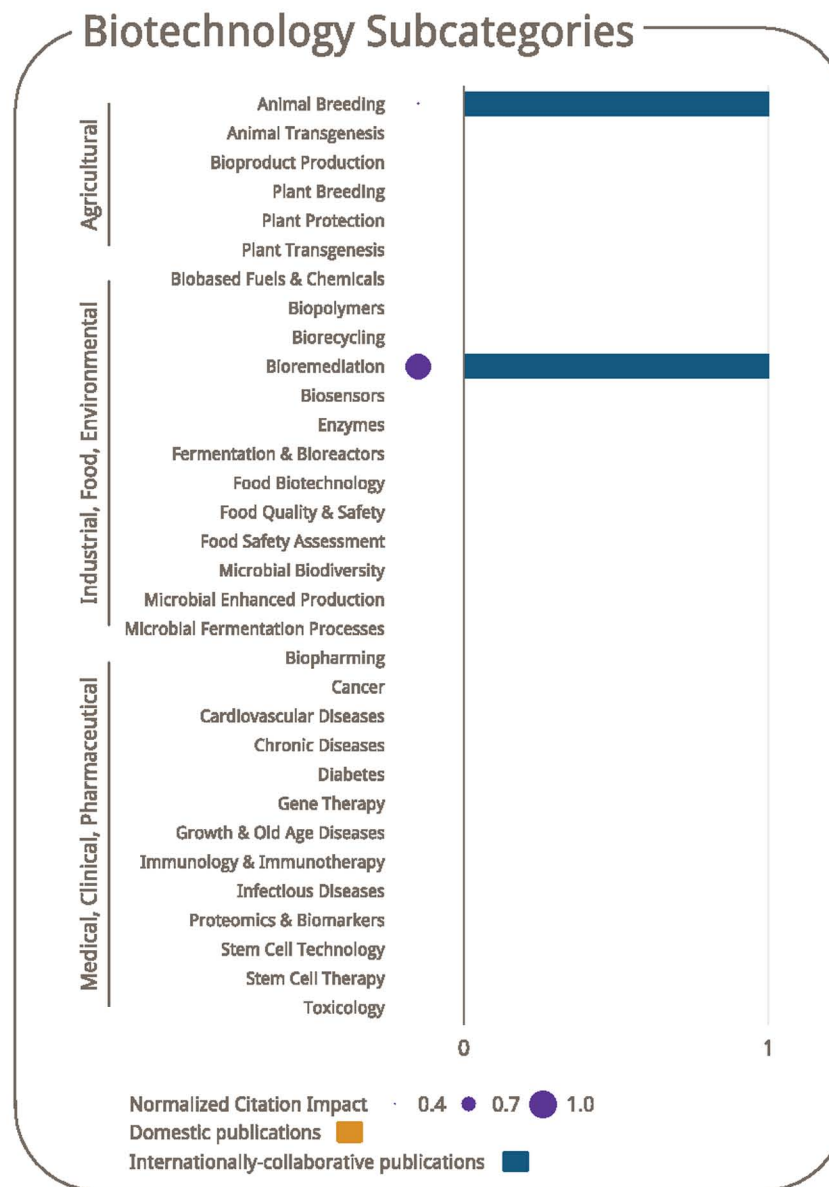
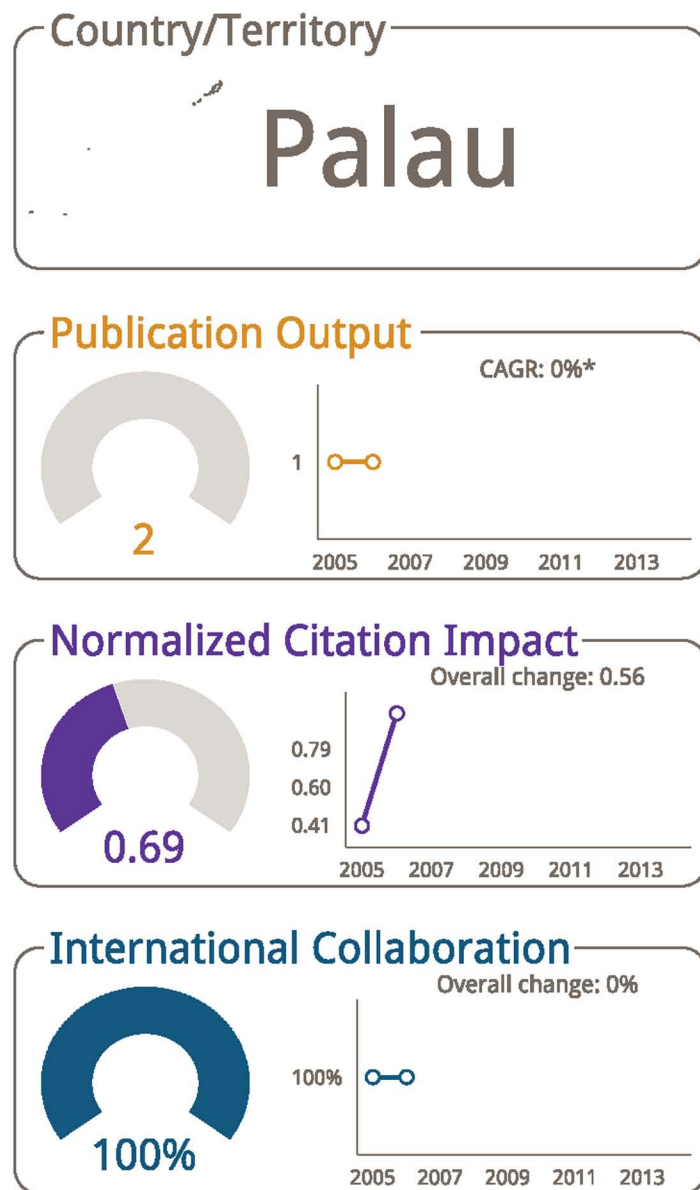


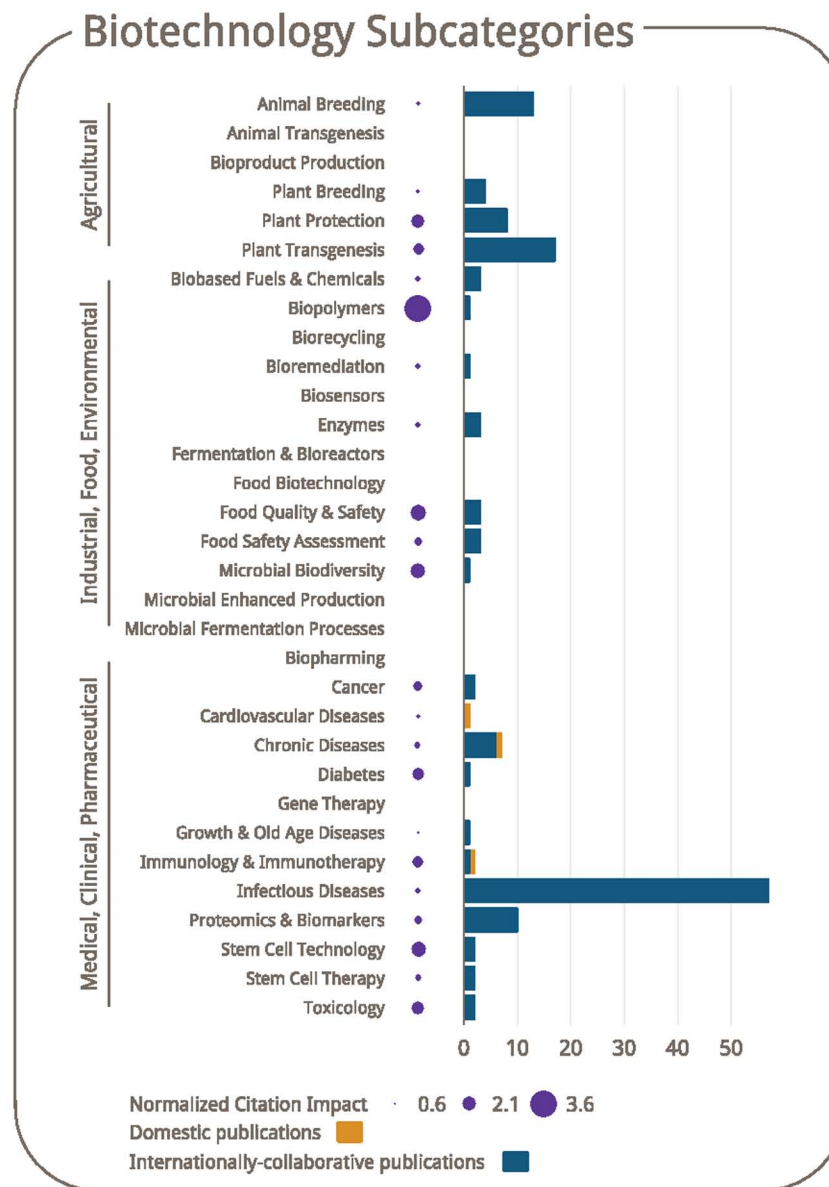
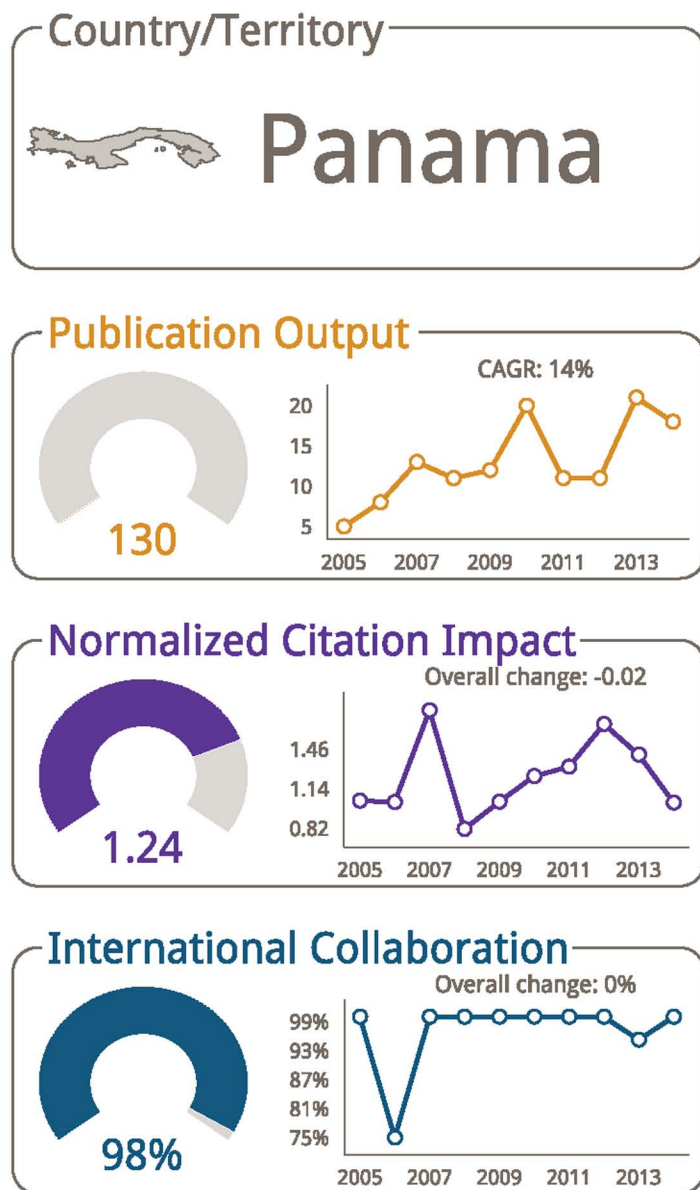












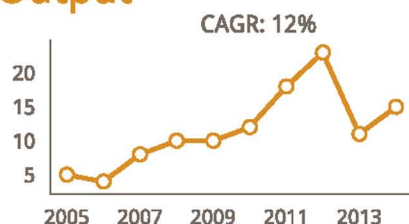


## Country/Territory

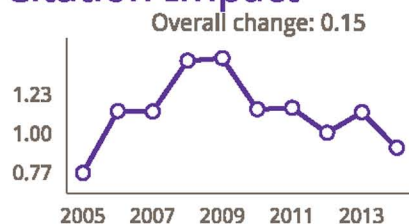


Papua New Guinea

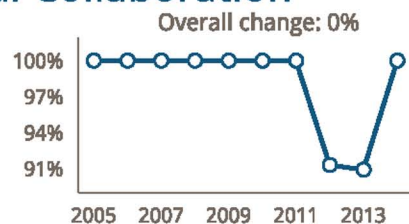
## Publication Output



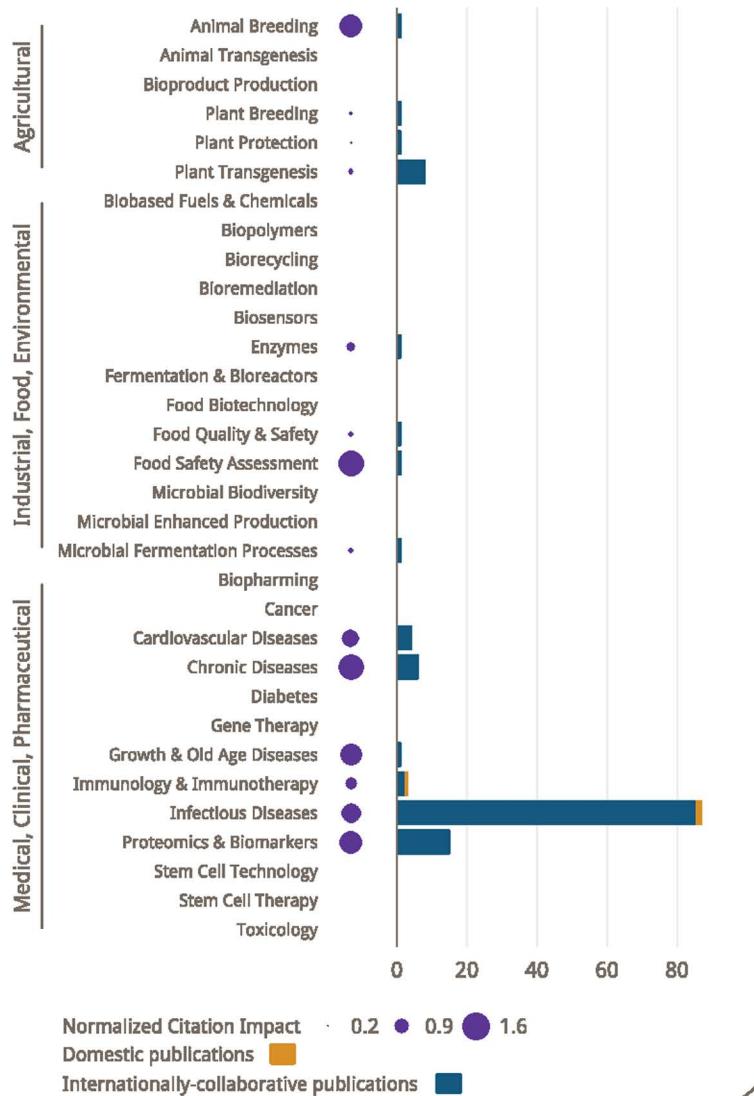
## Normalized Citation Impact



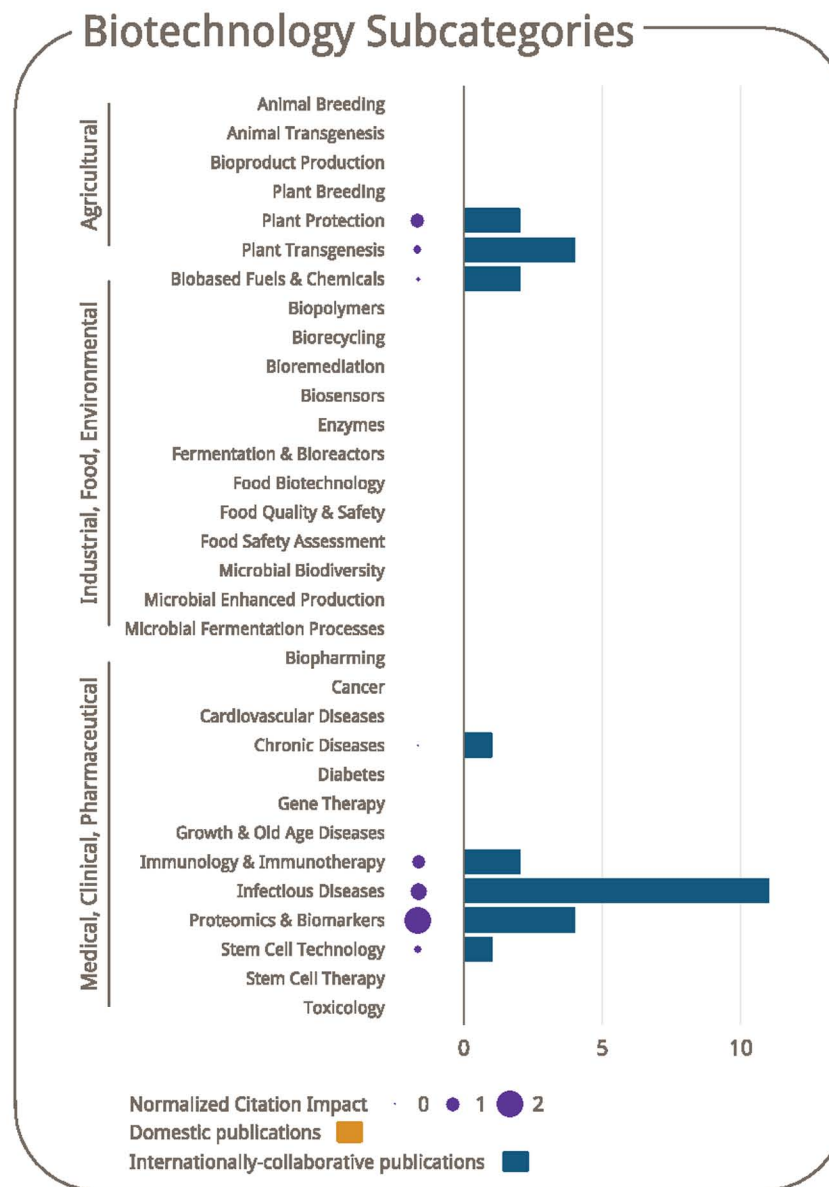
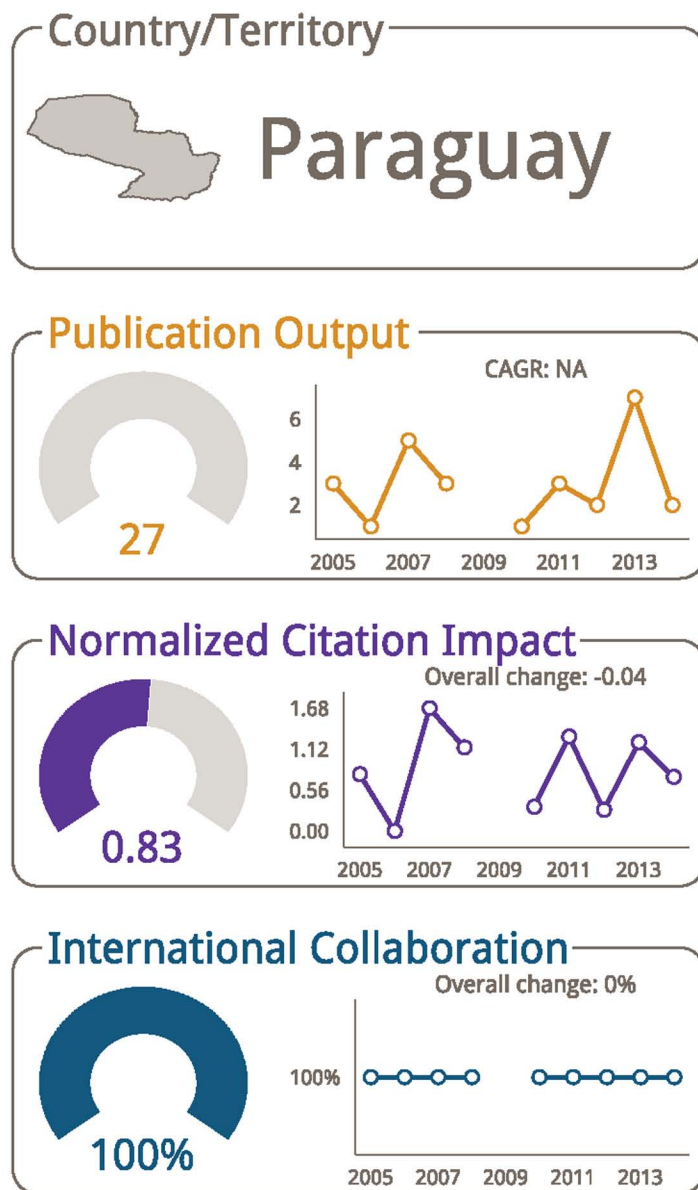
## International Collaboration

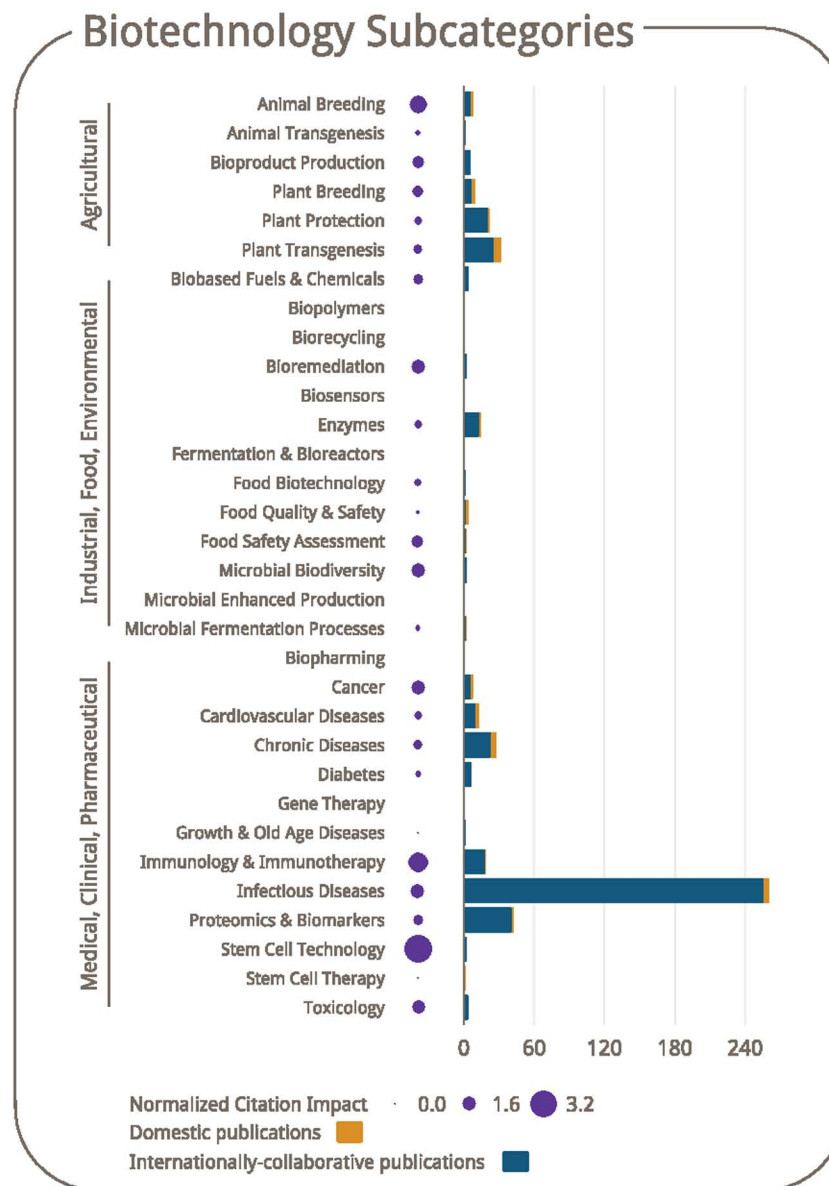
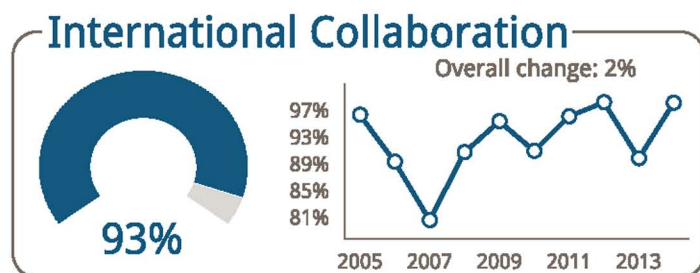
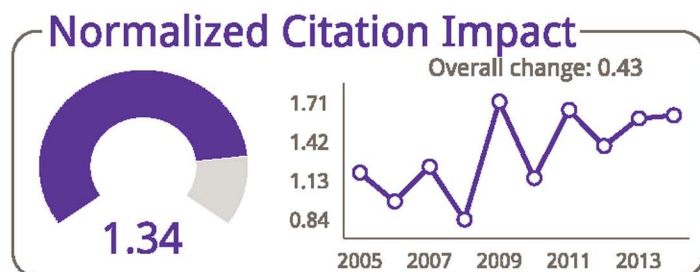
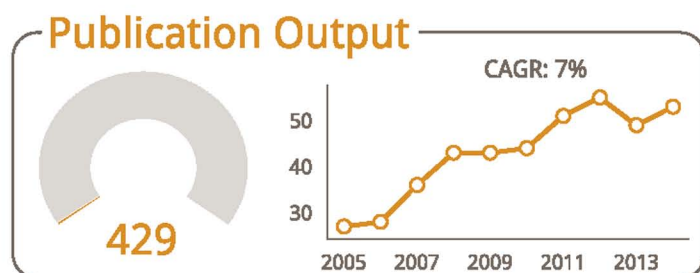


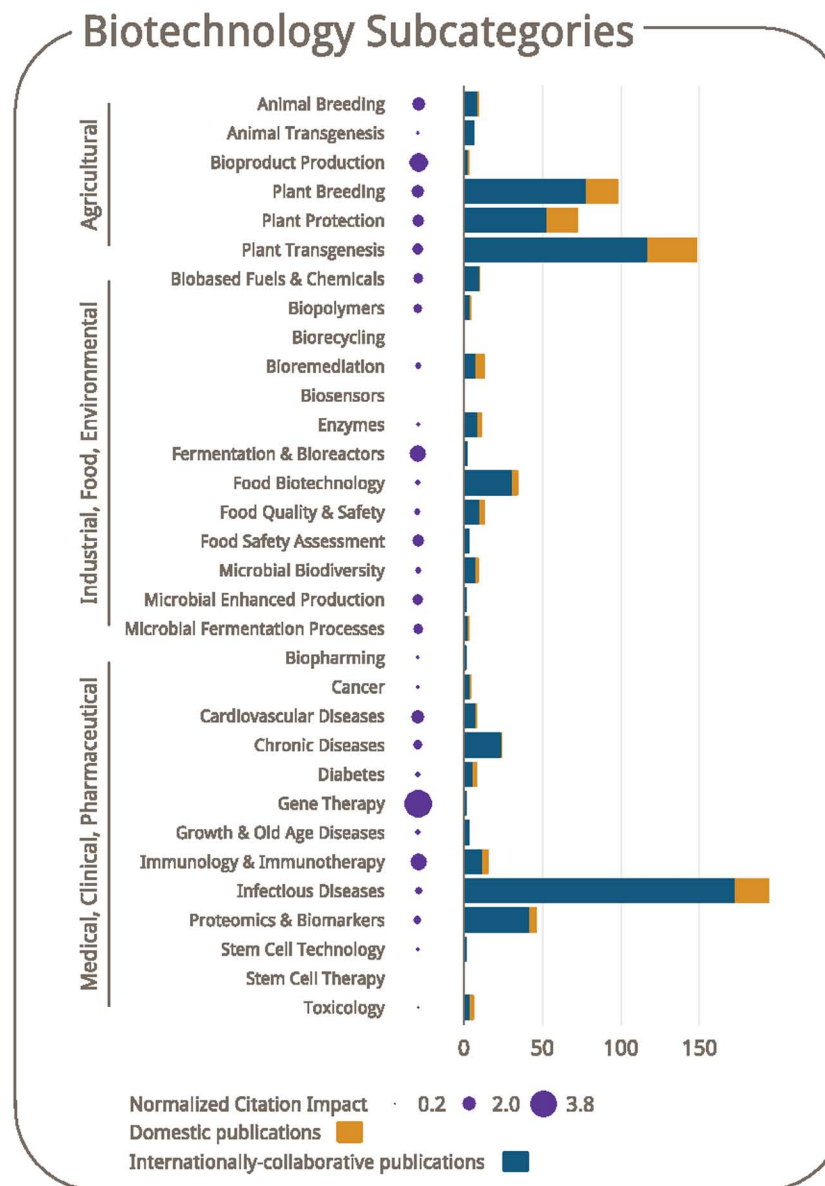
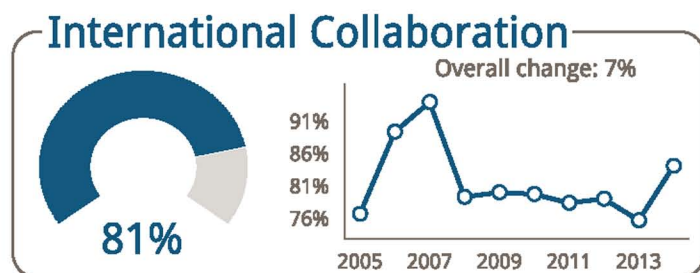
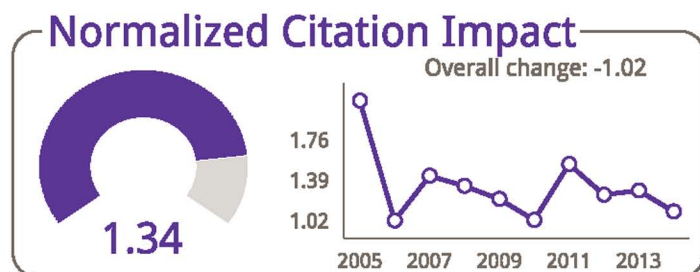
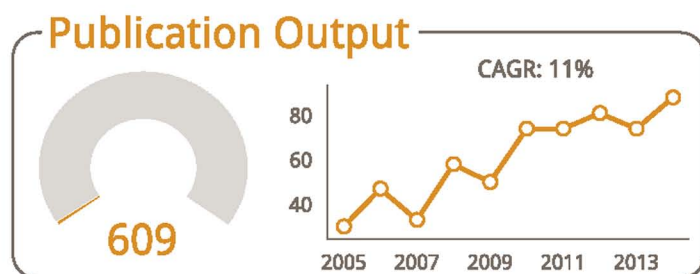
## Biotechnology Subcategories

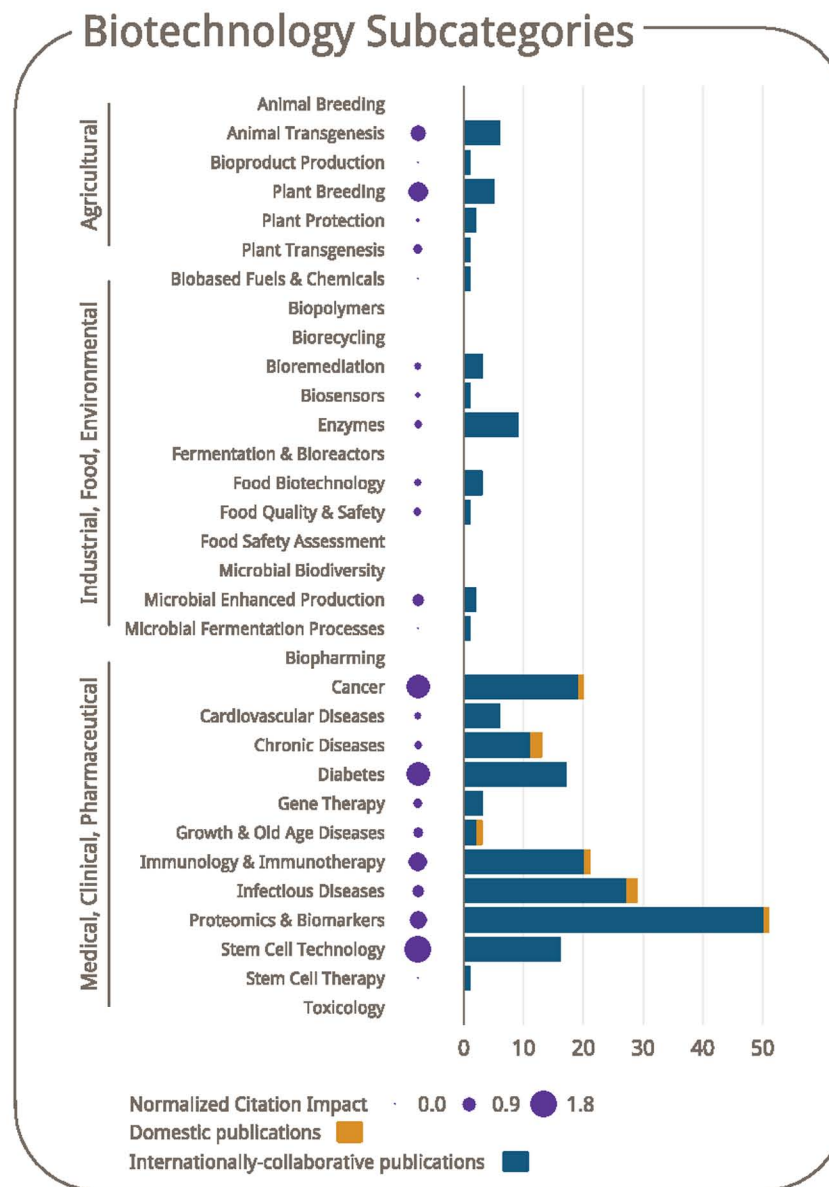
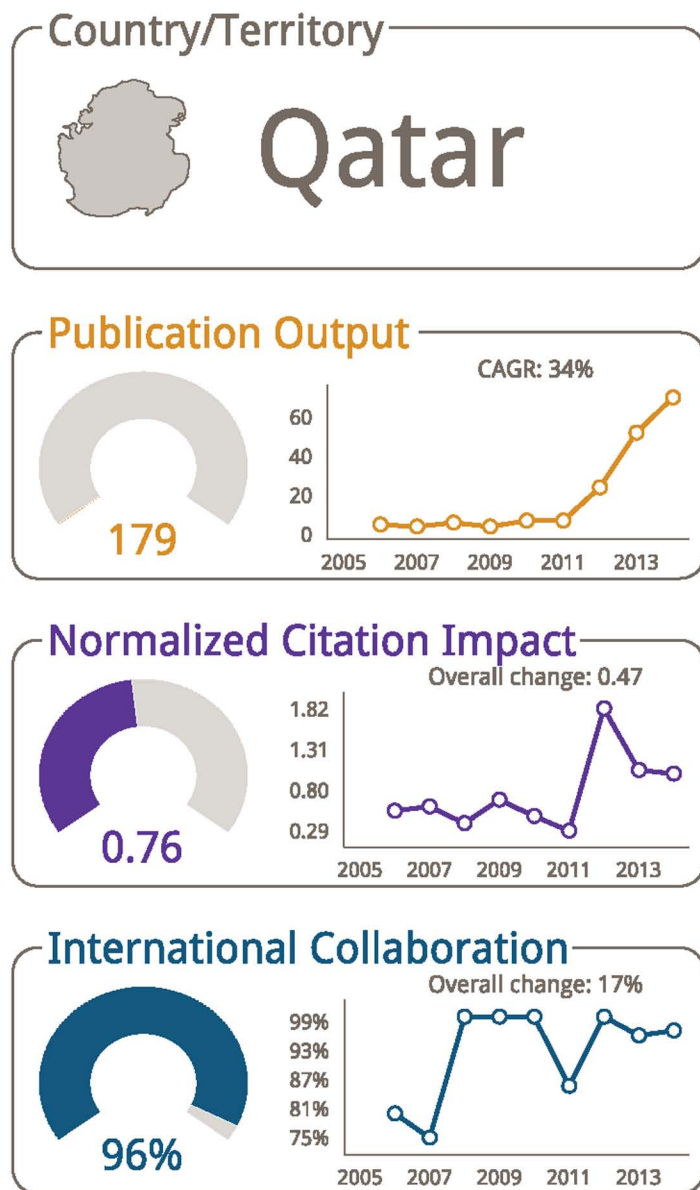


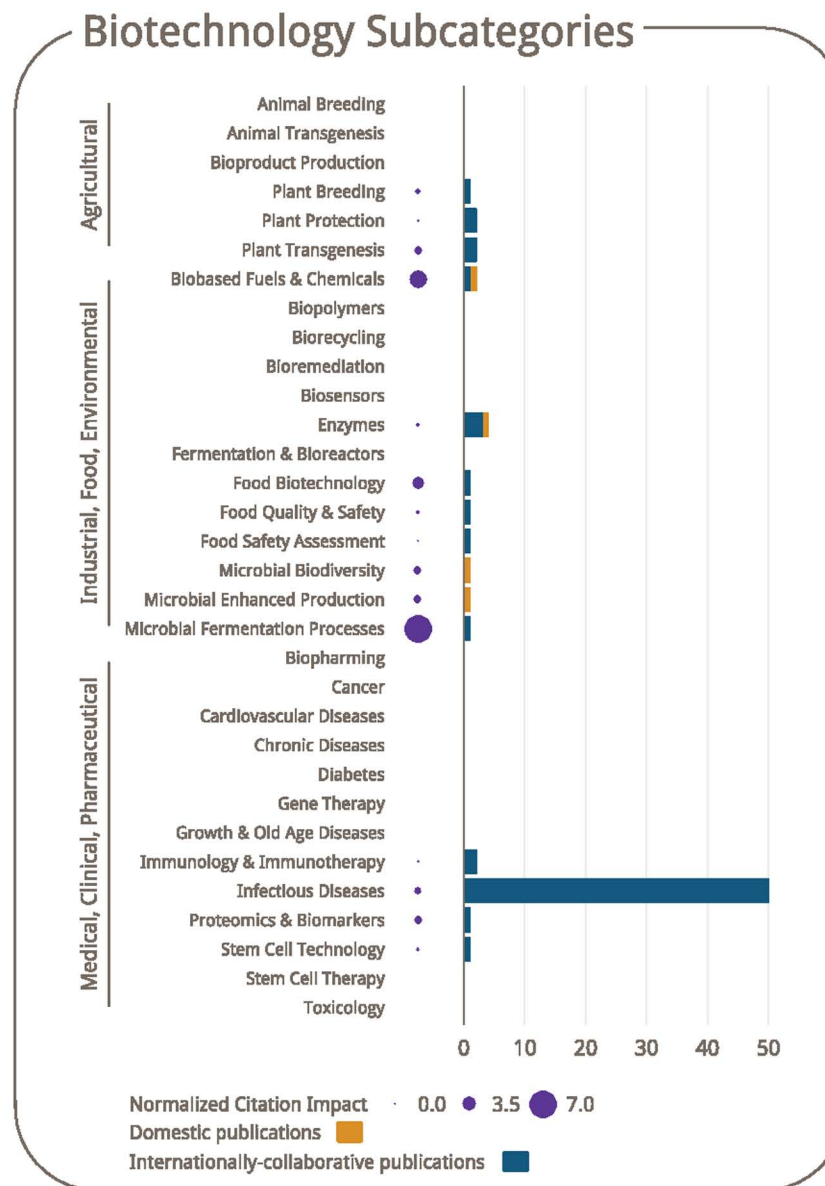
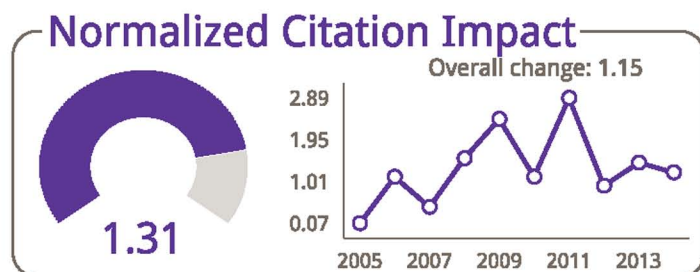
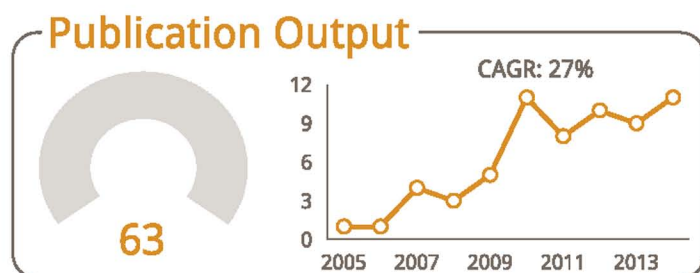




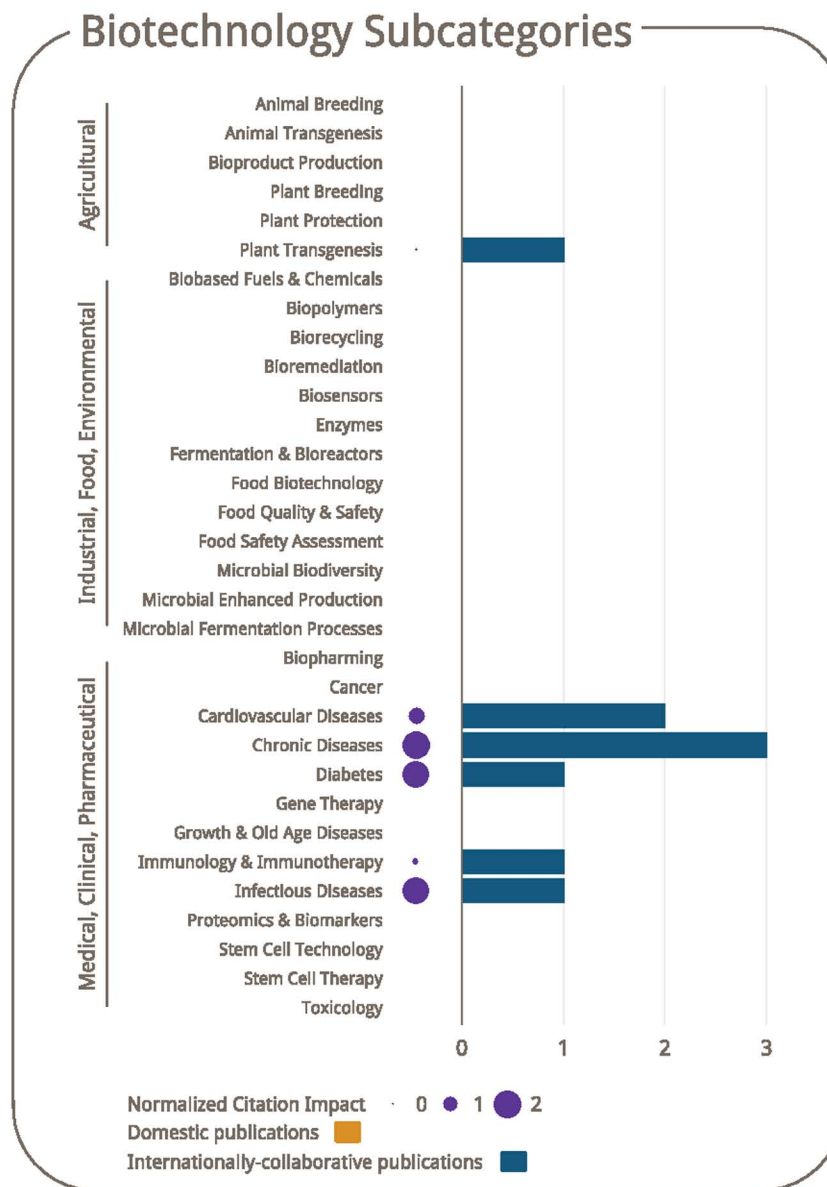
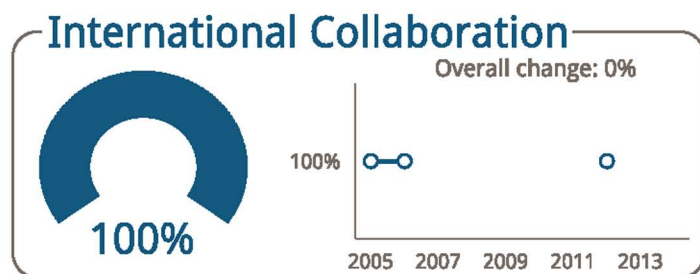
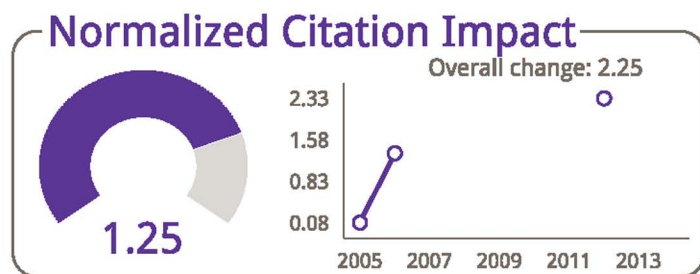
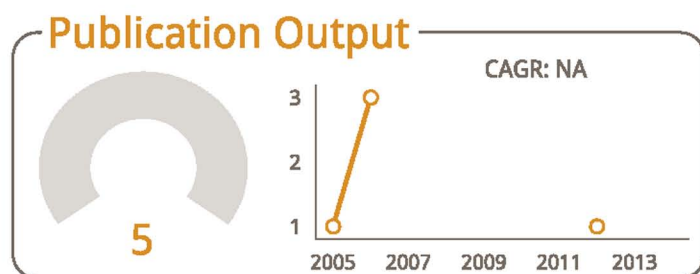




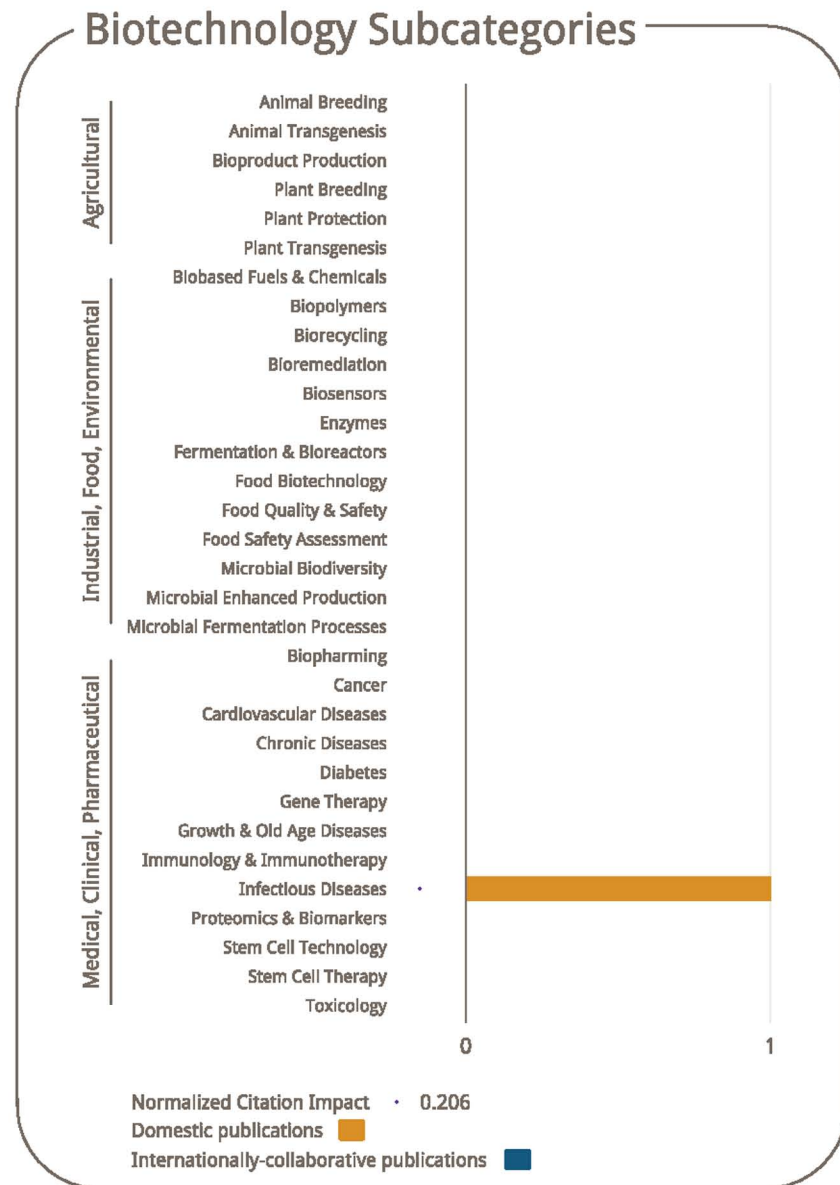
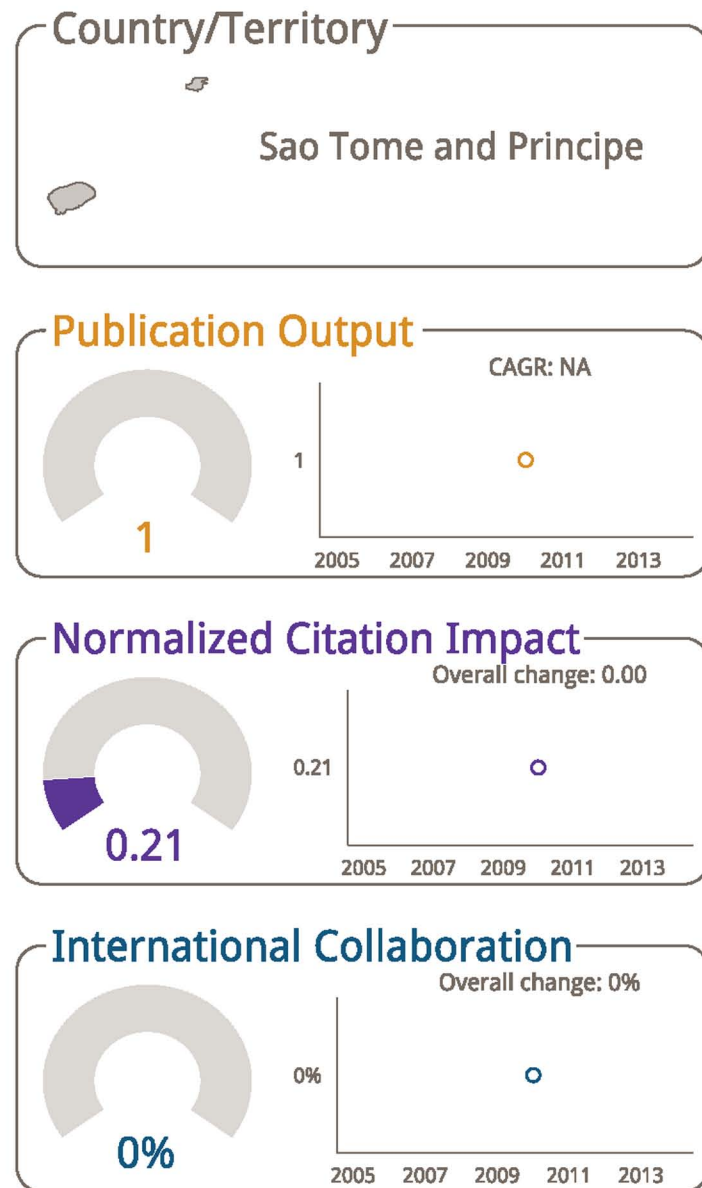


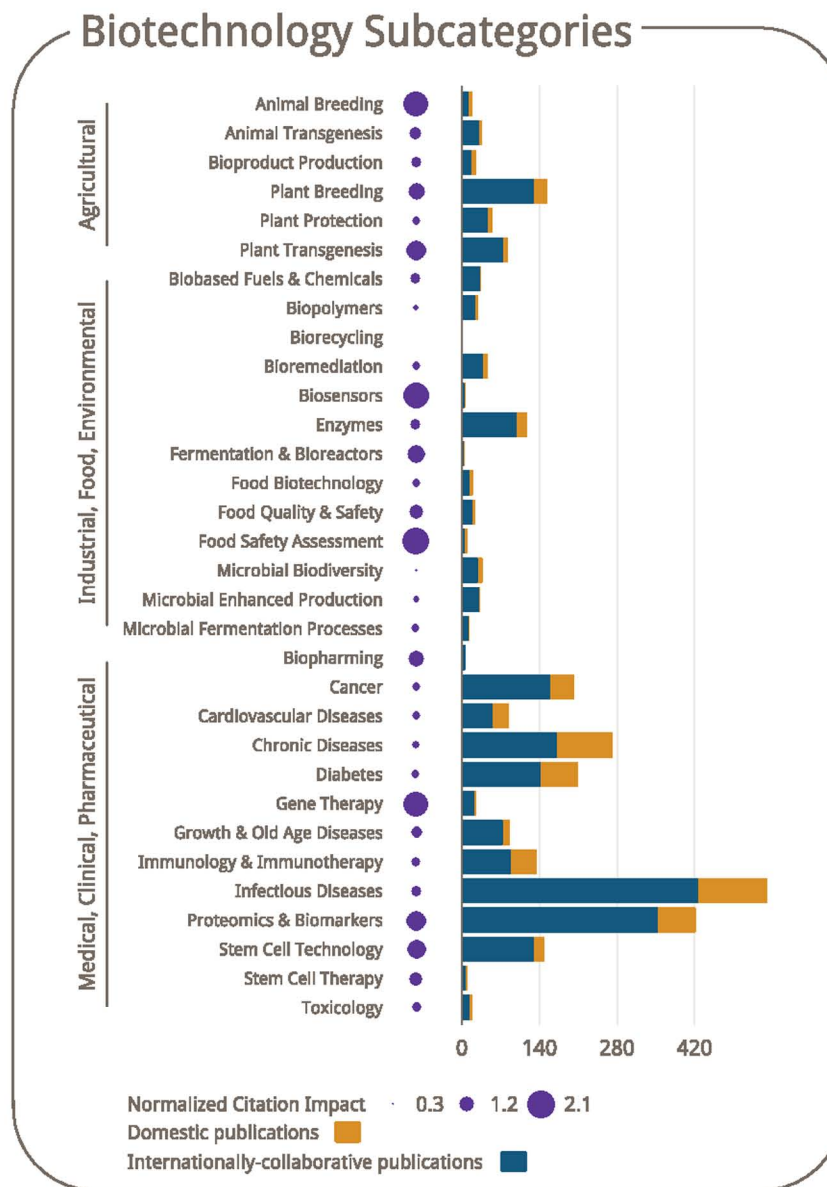
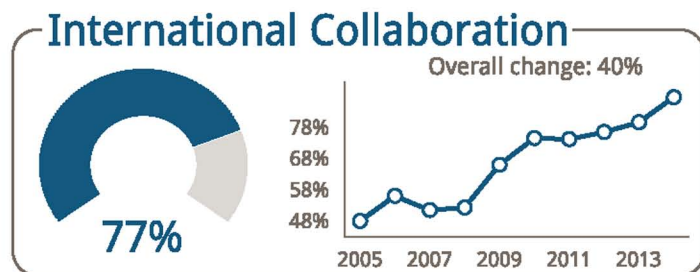
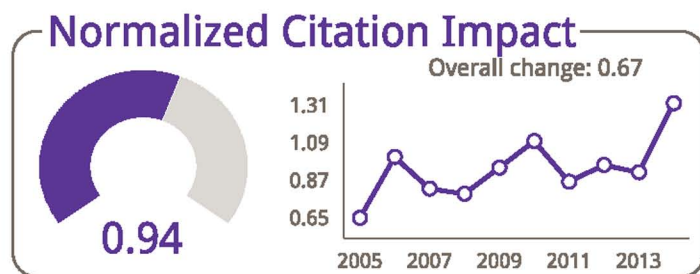
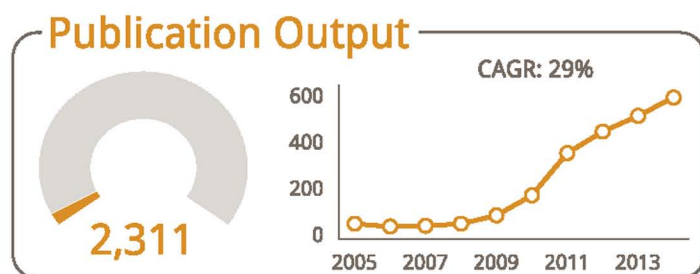


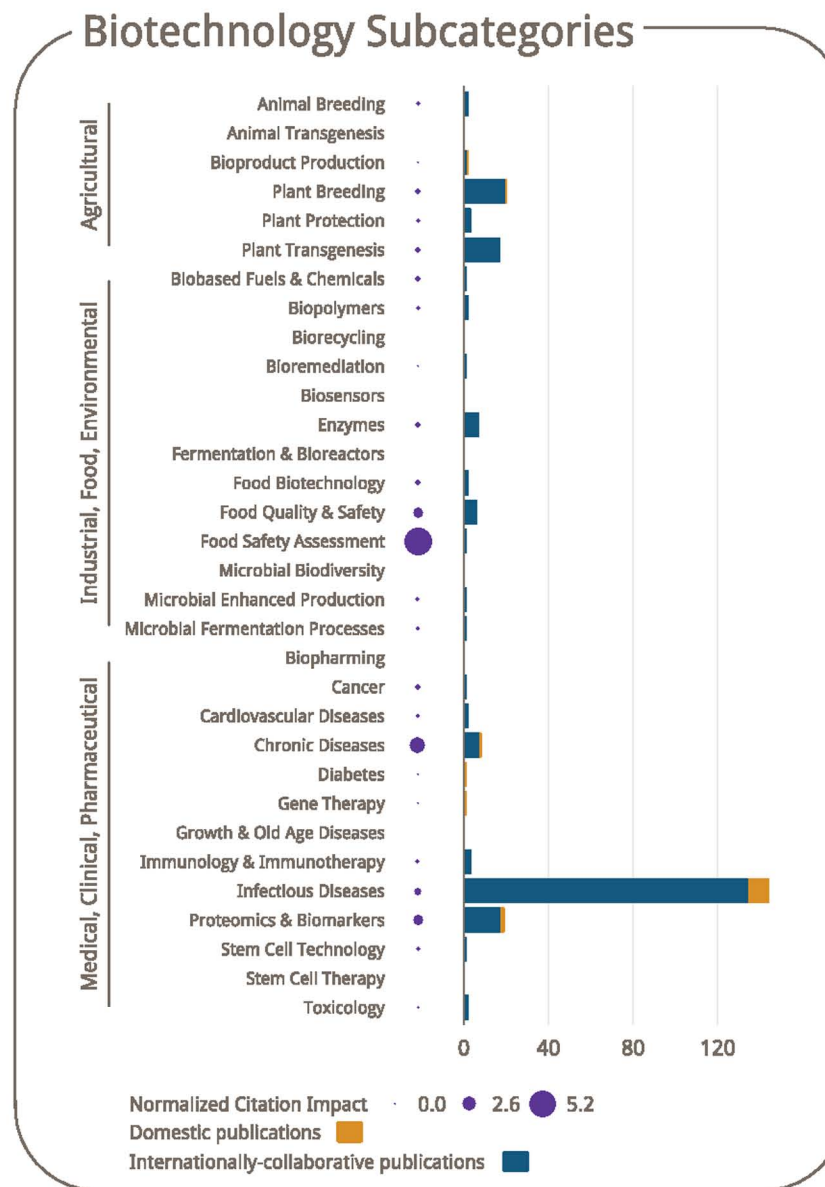
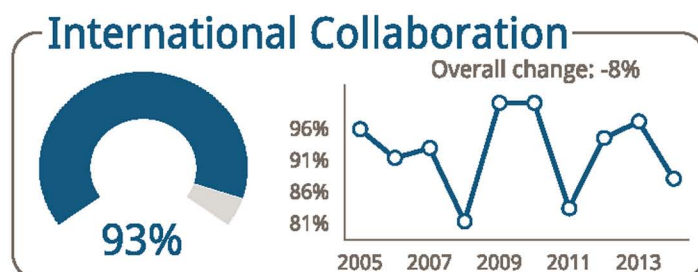
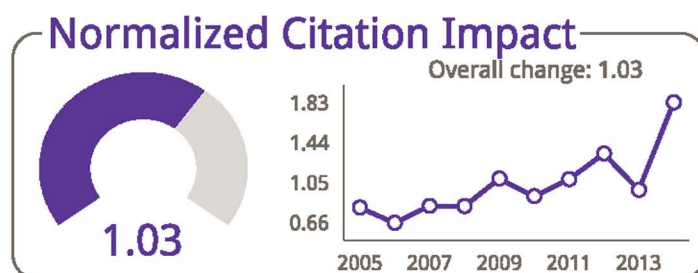
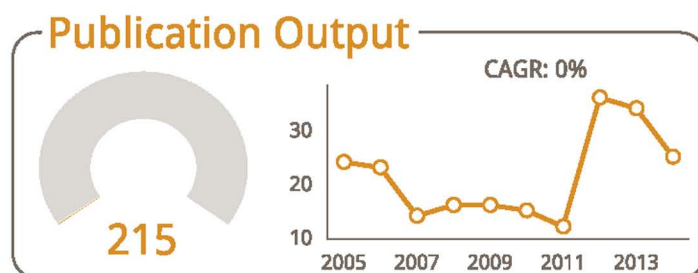


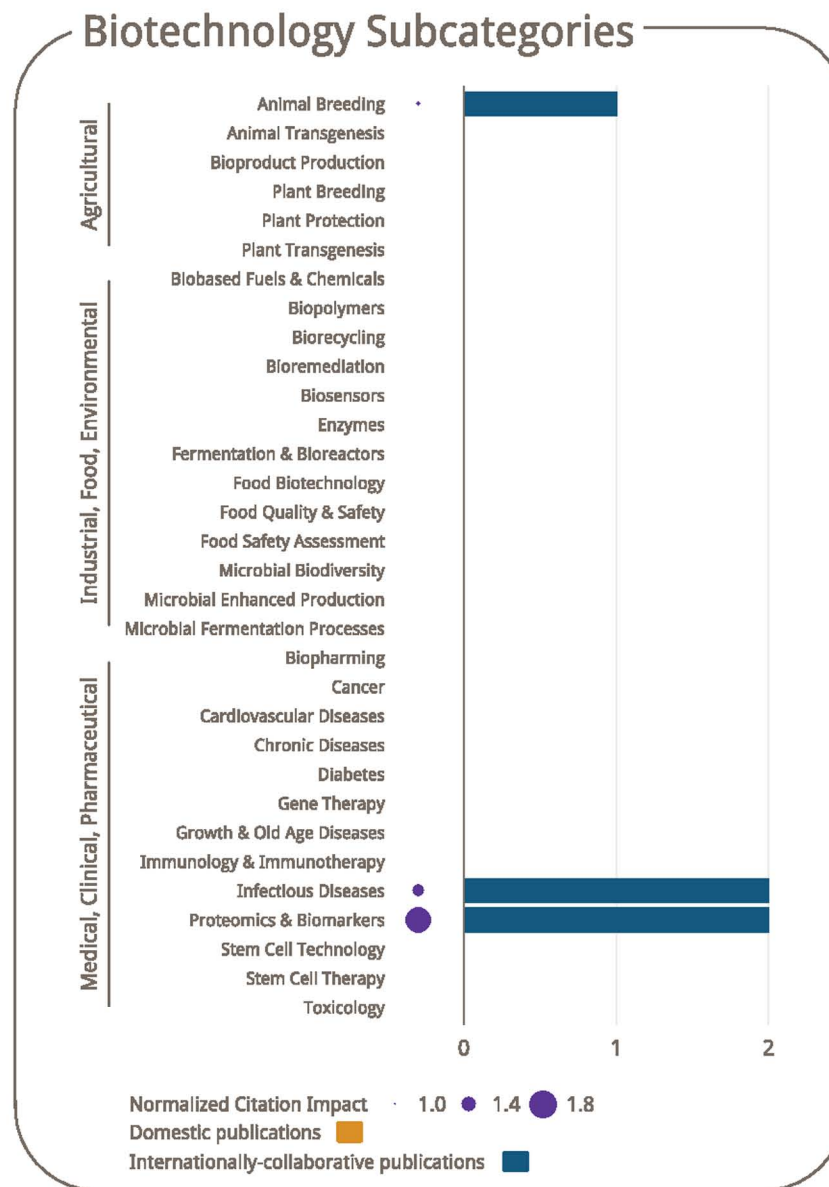
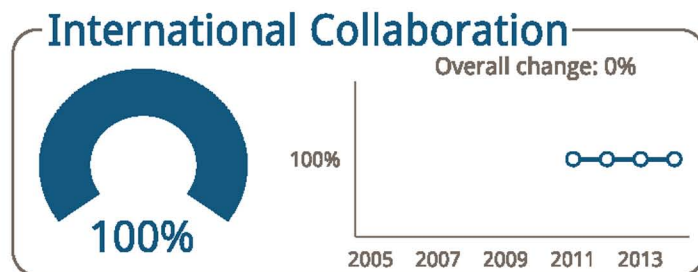
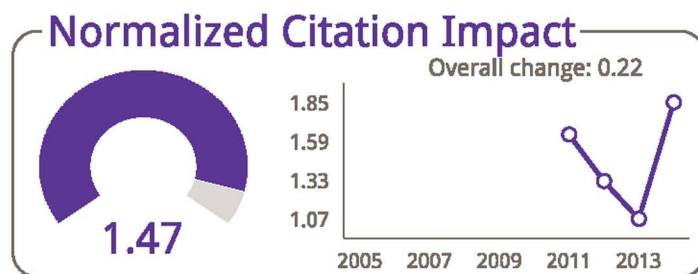
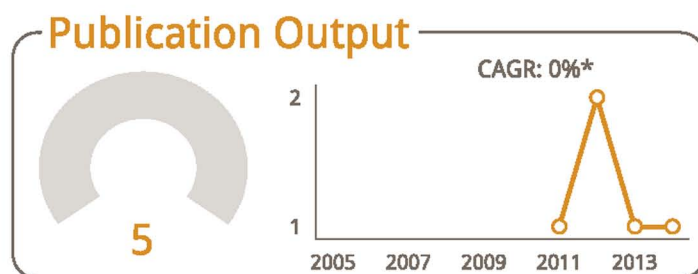


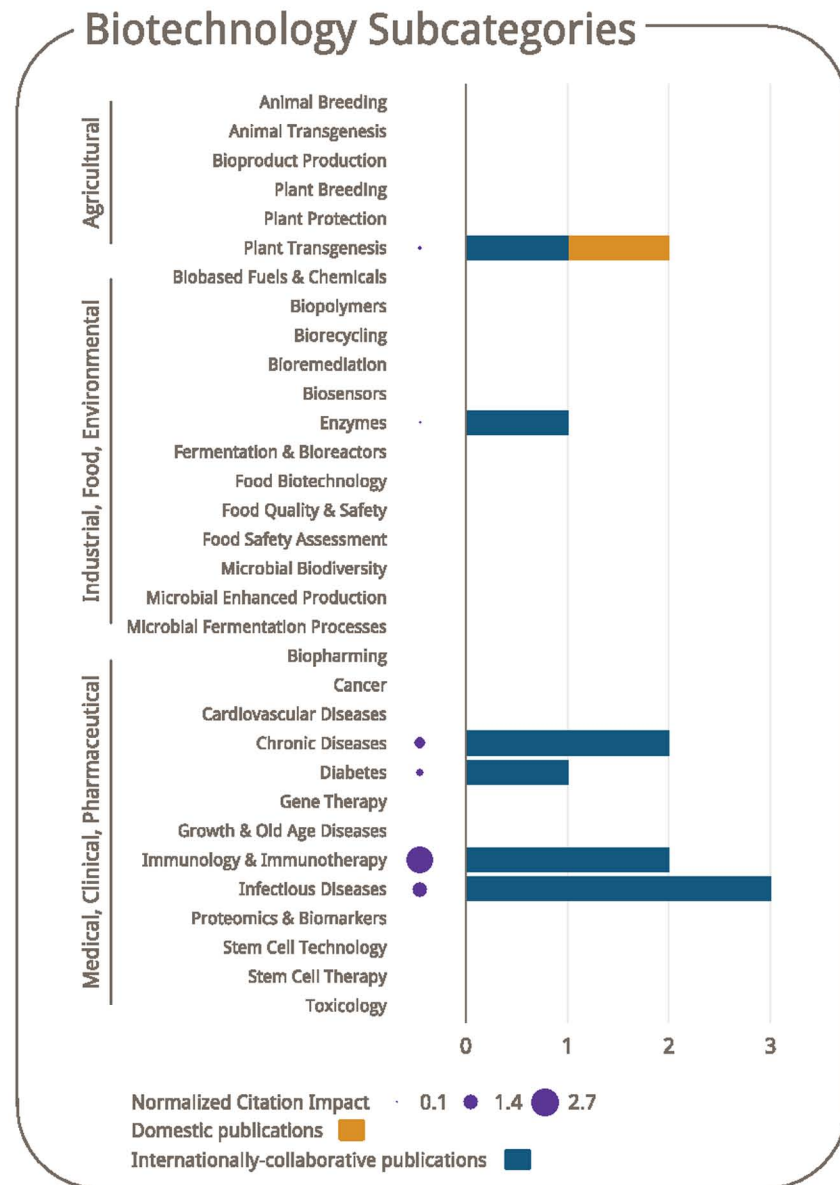
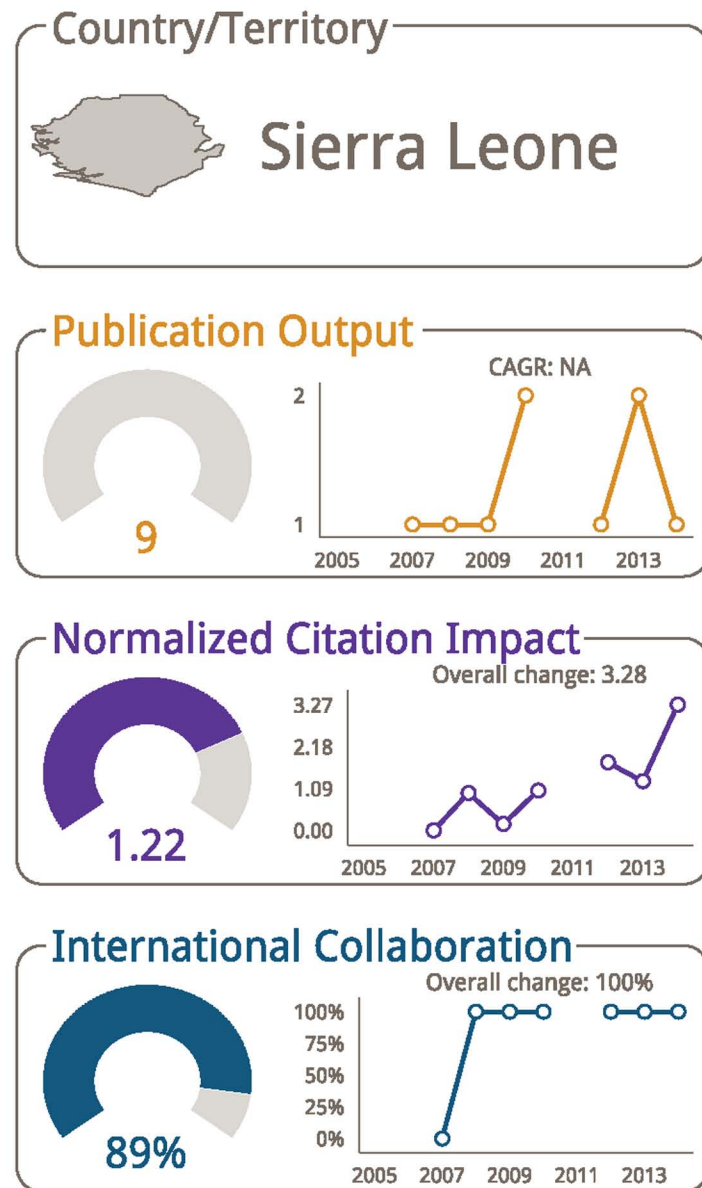




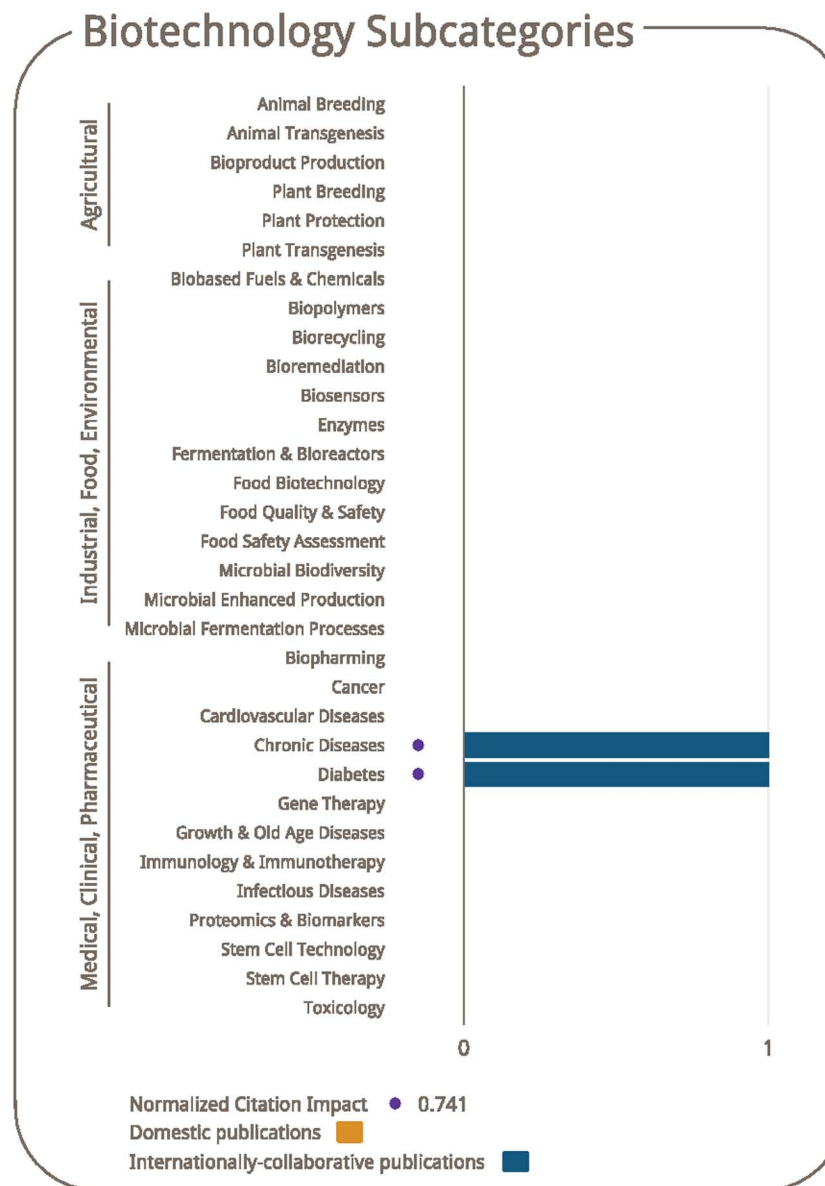
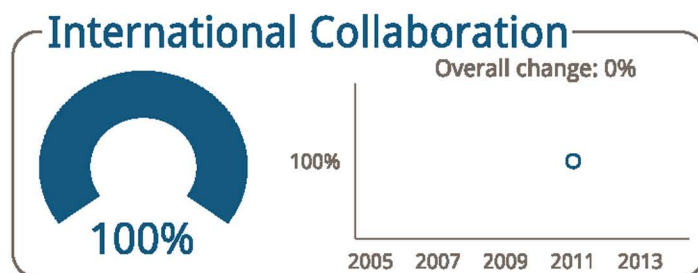
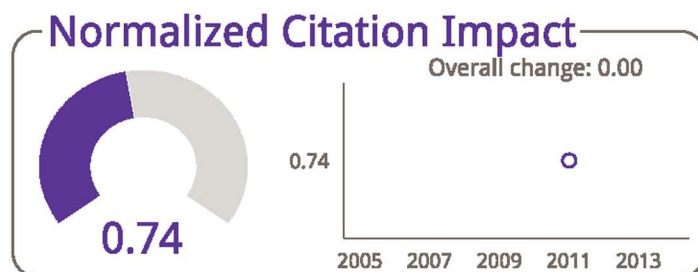
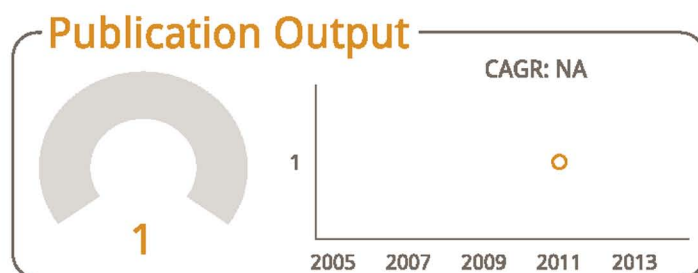
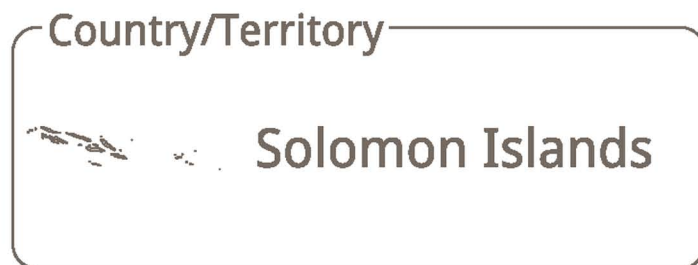




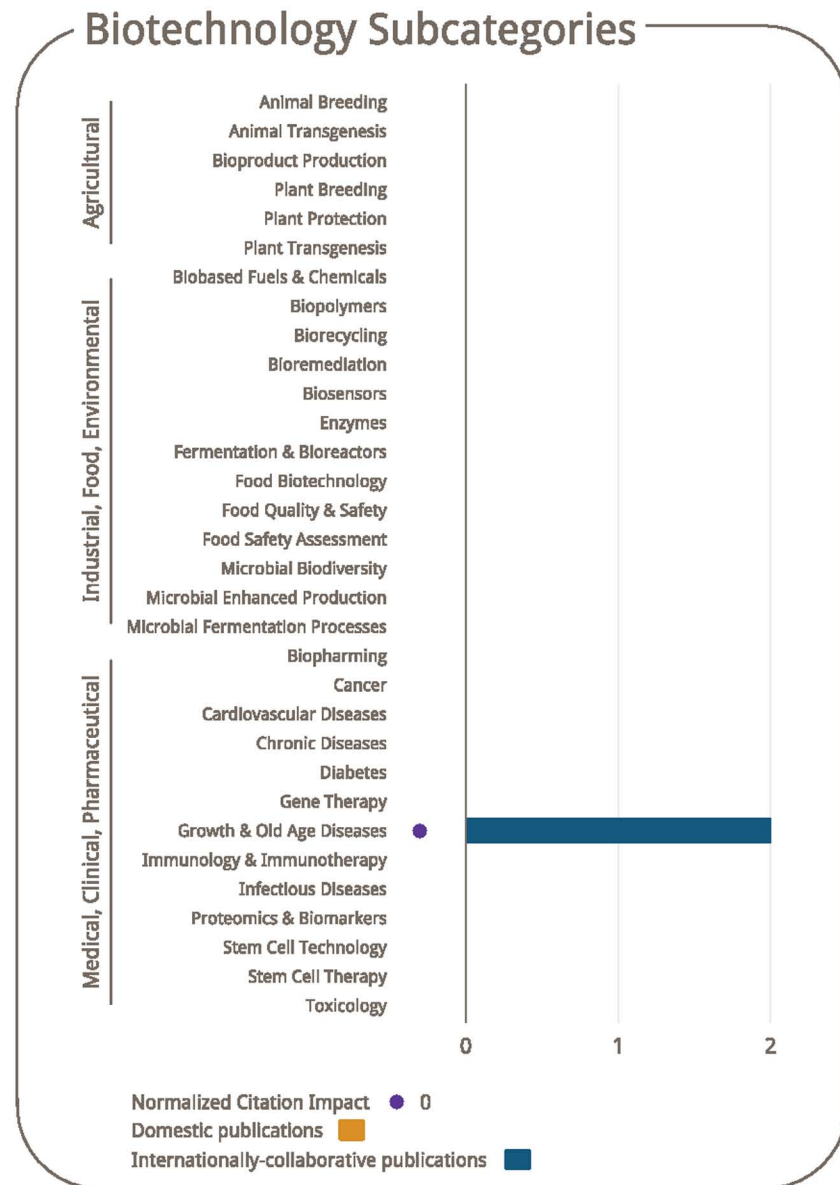
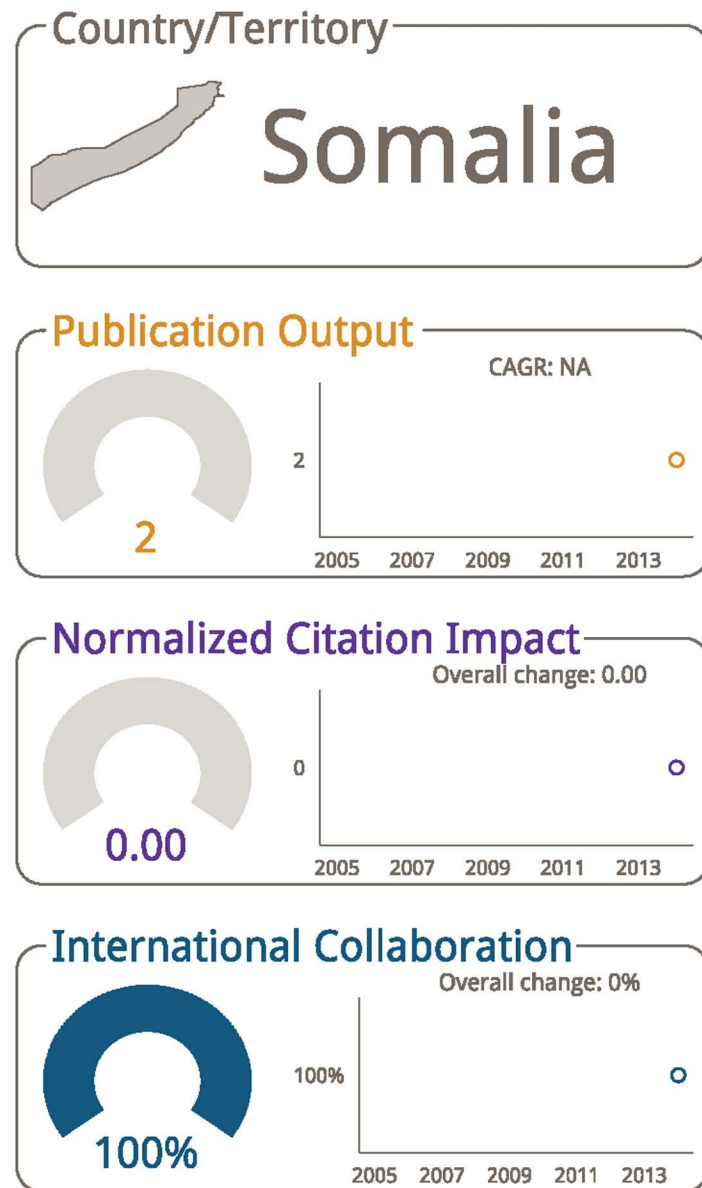


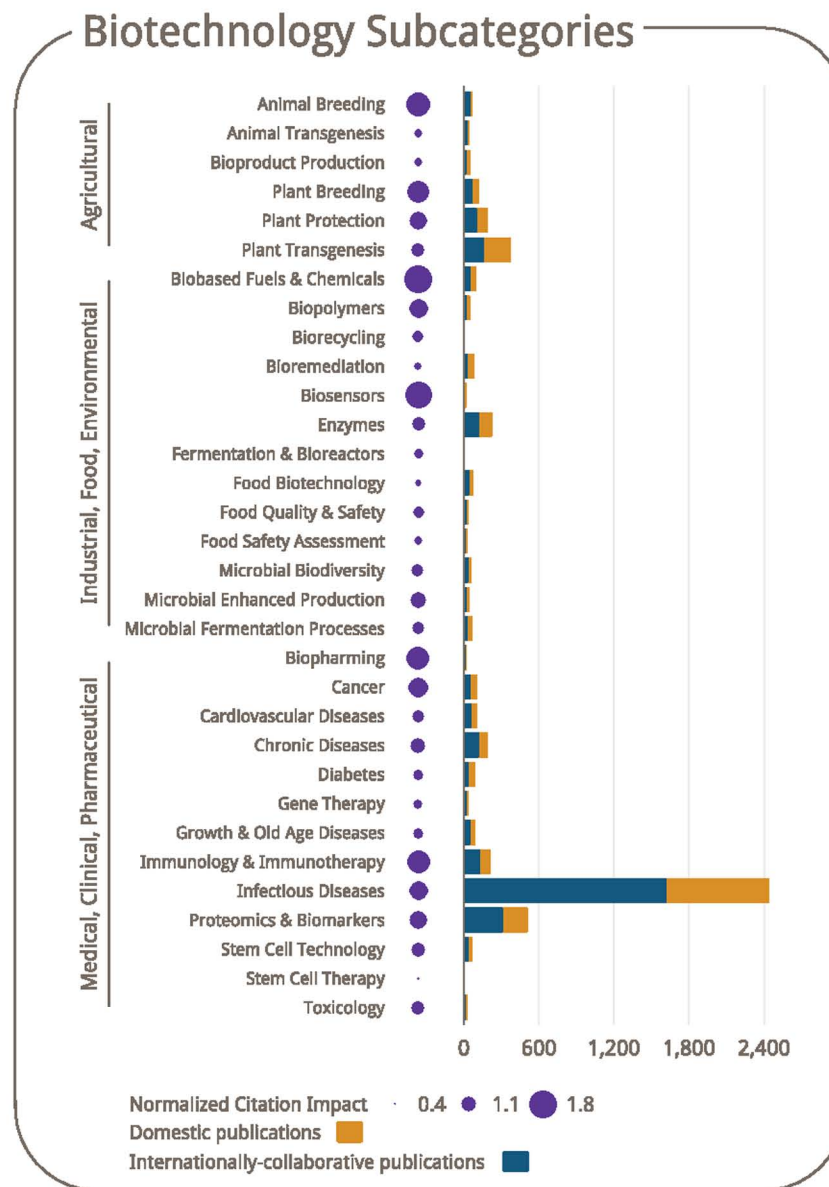
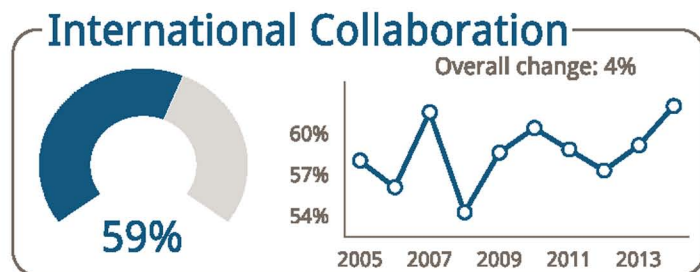
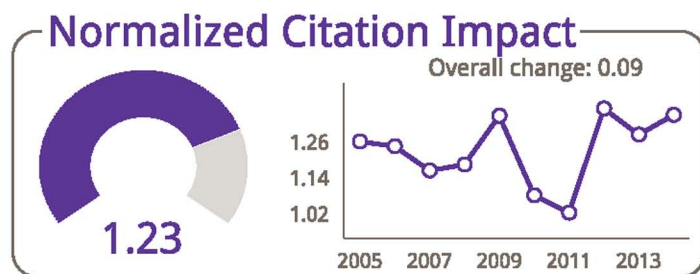
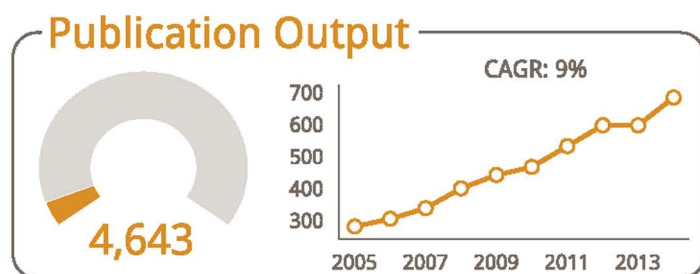


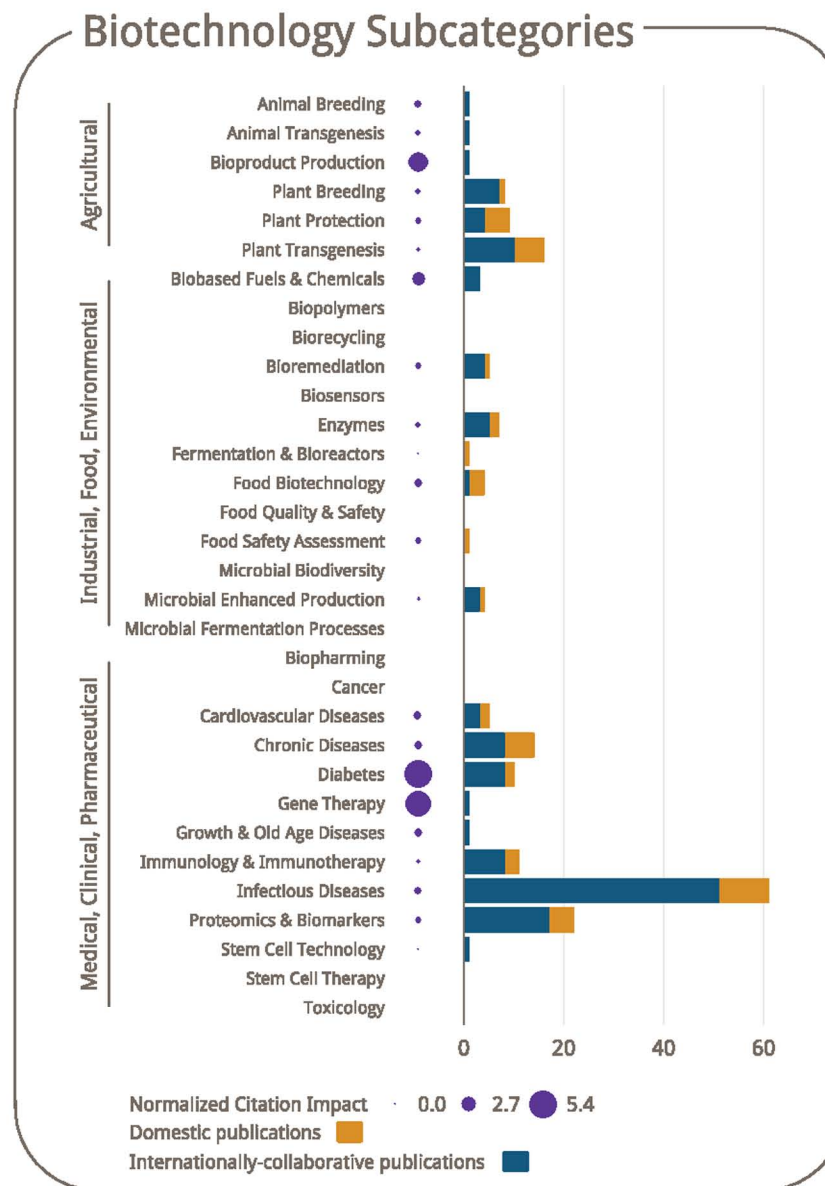
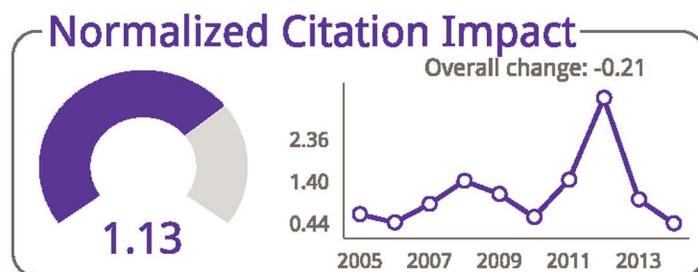
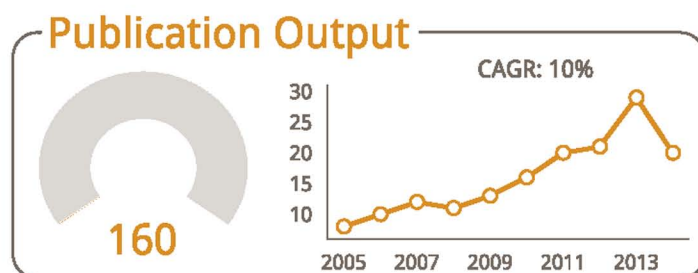


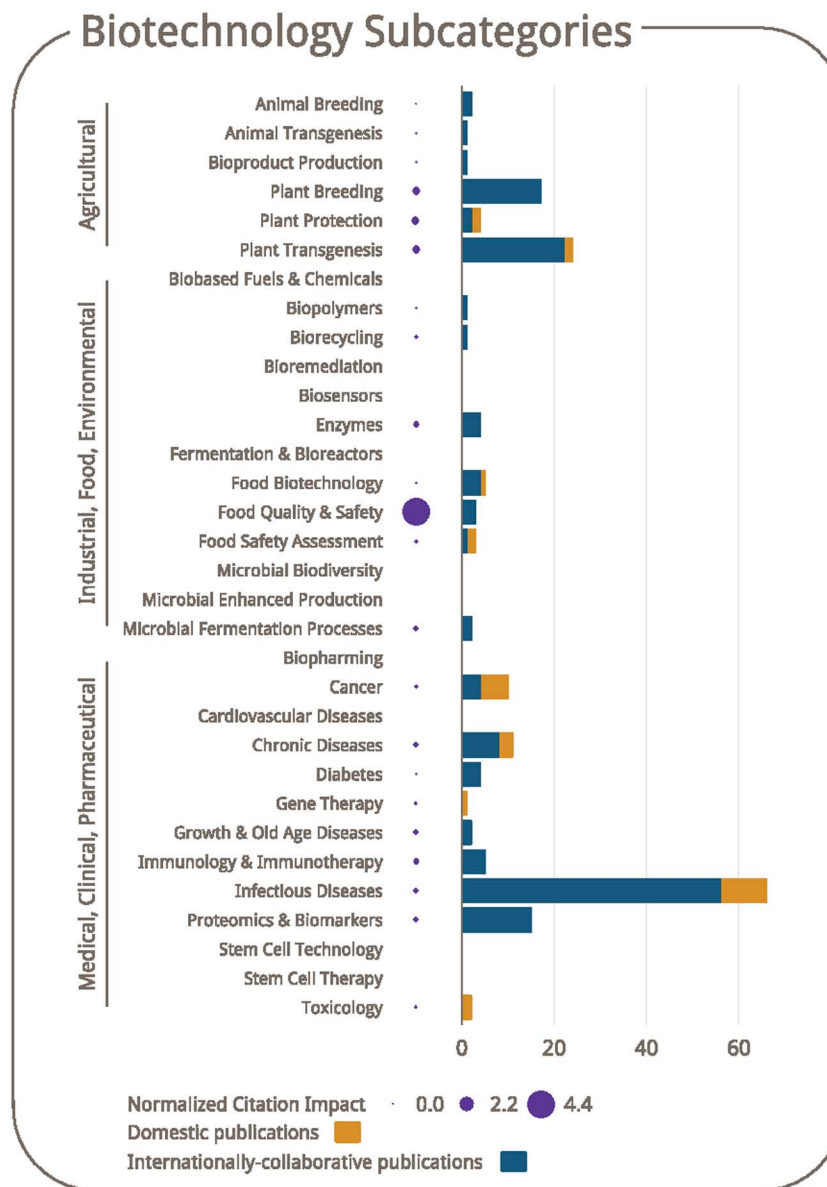
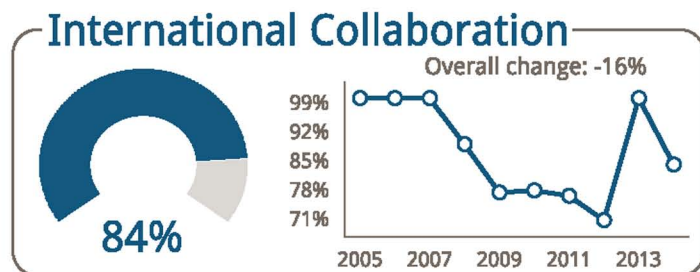
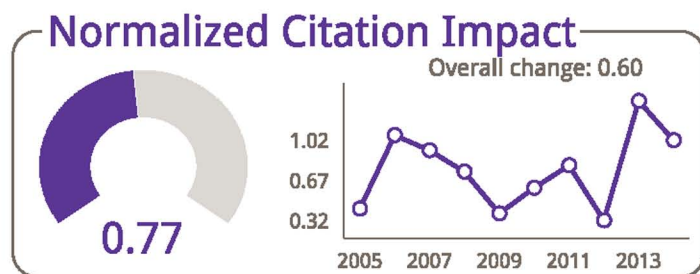
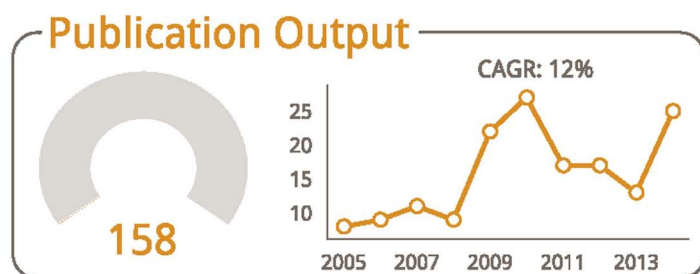


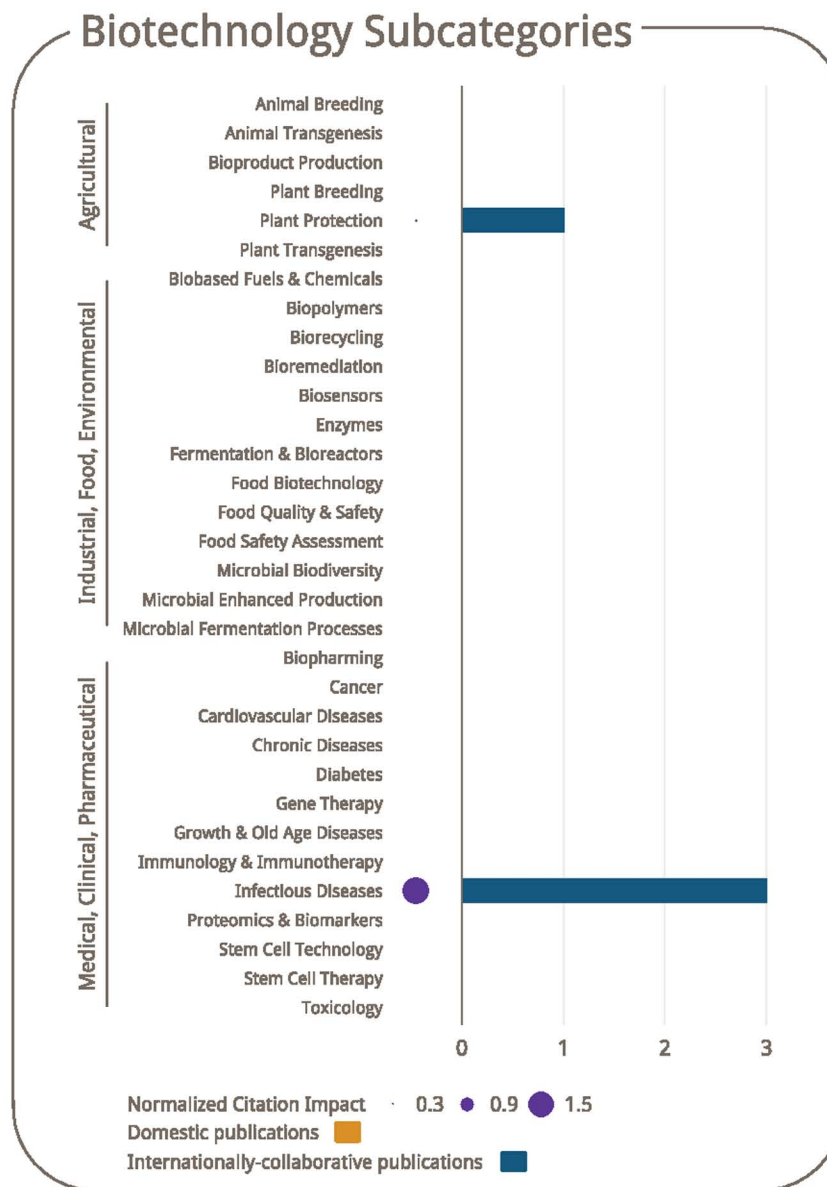
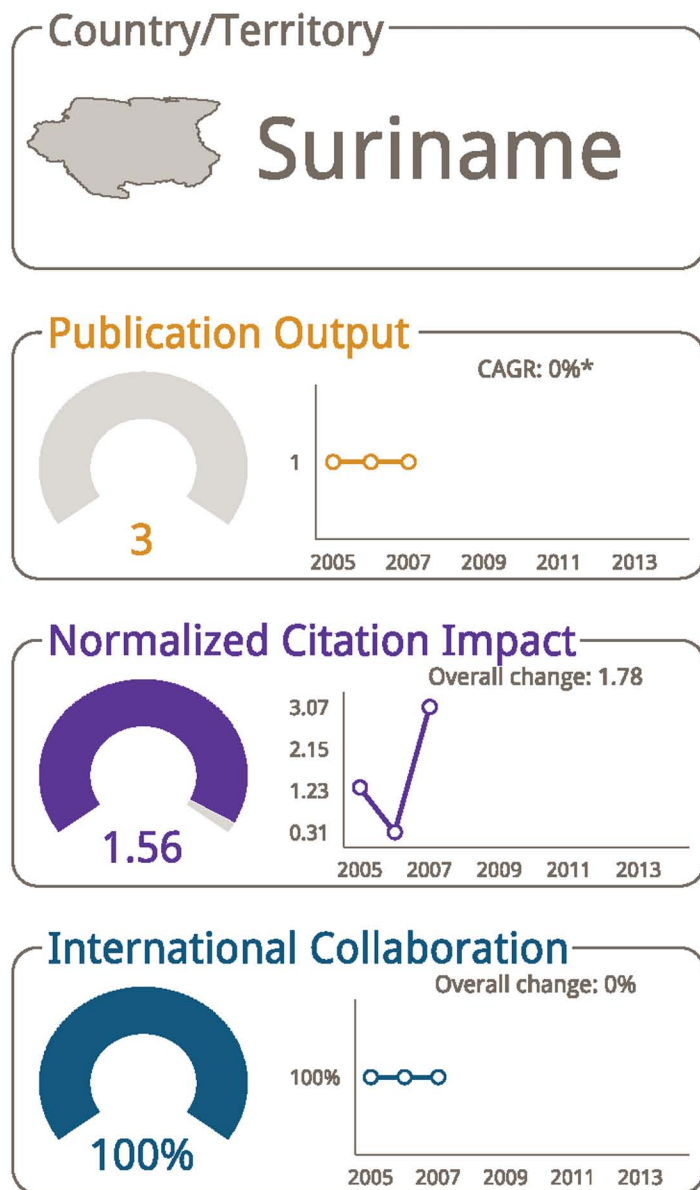




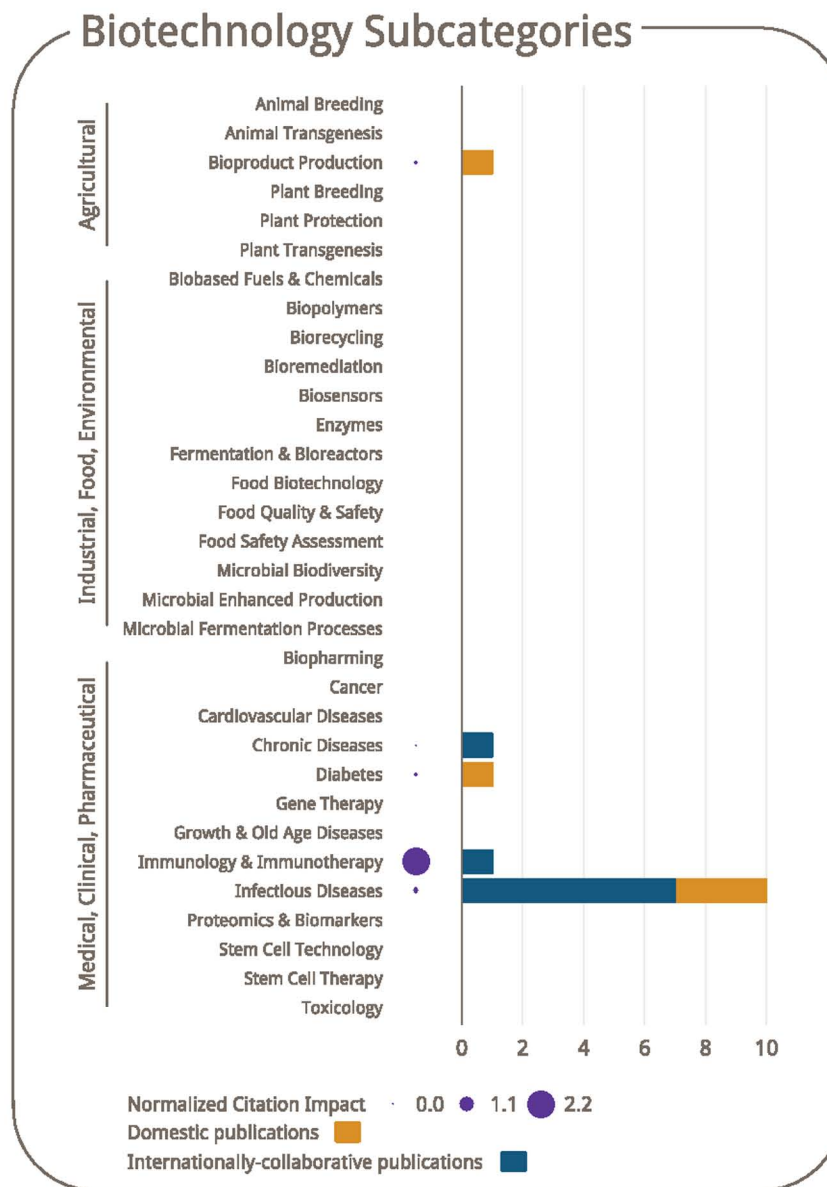
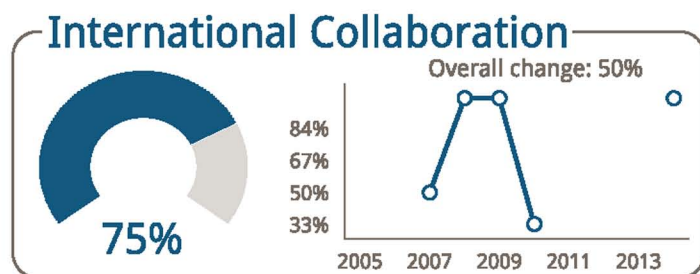
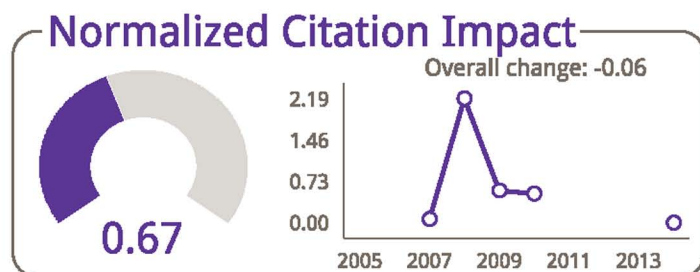
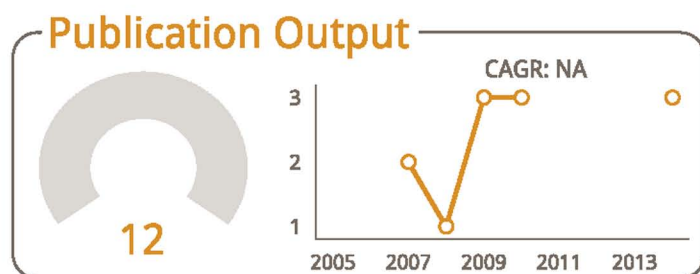




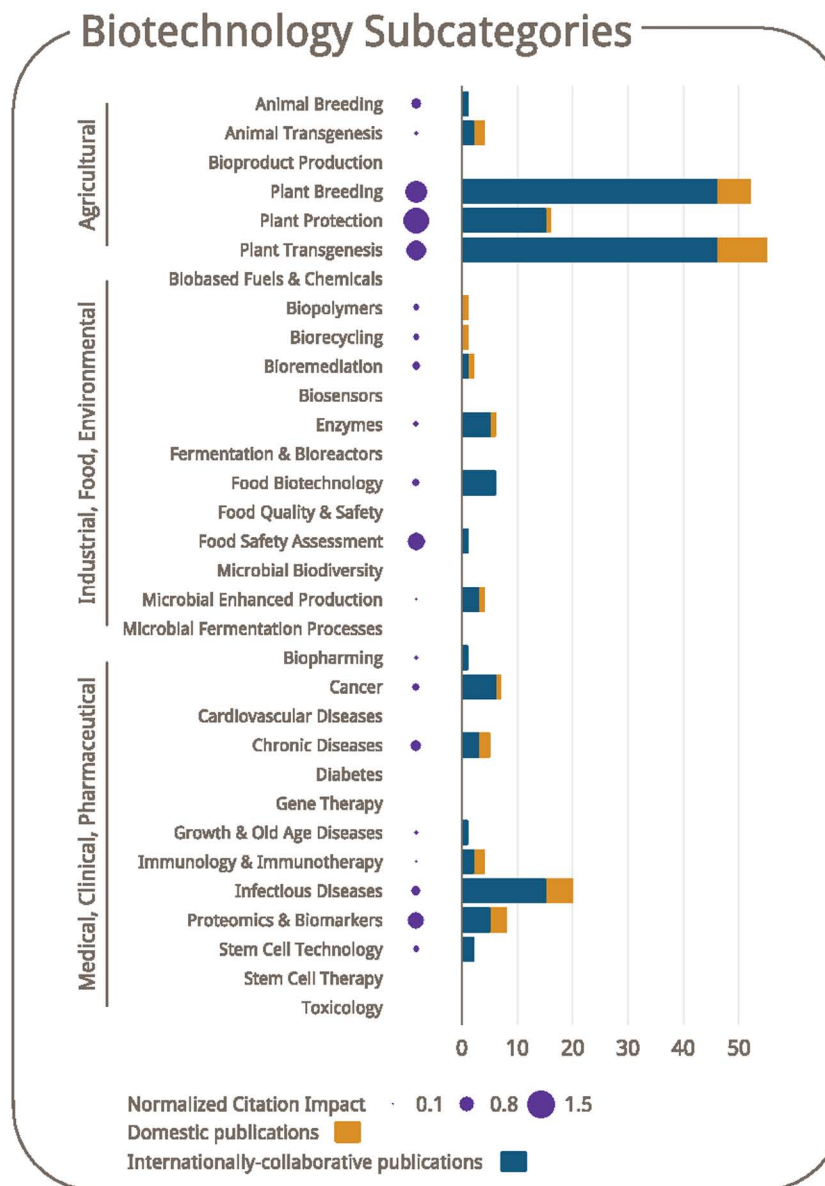
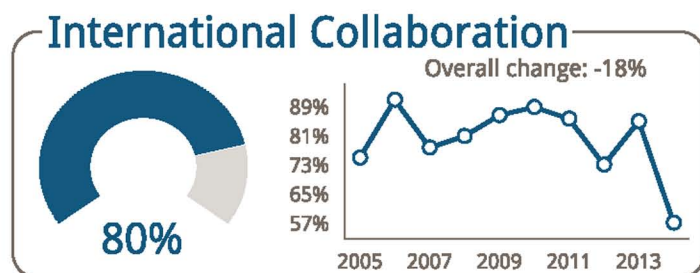
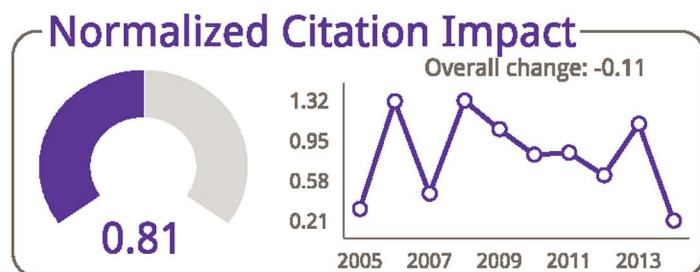
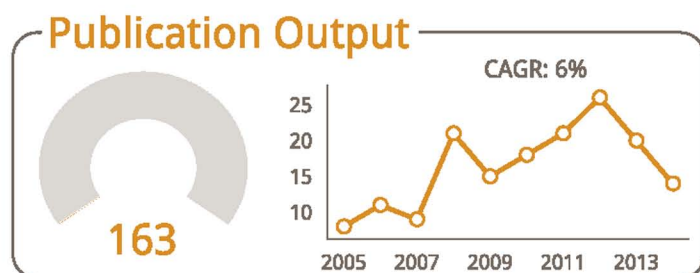


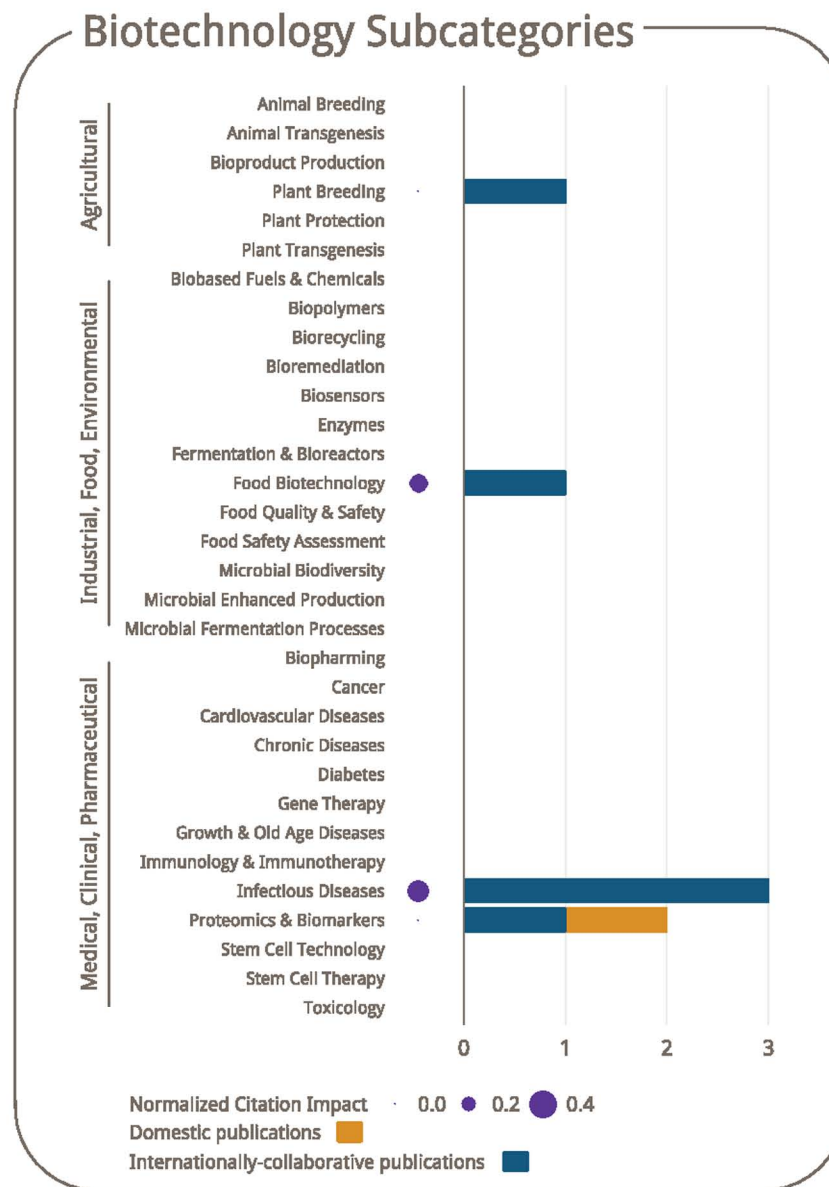
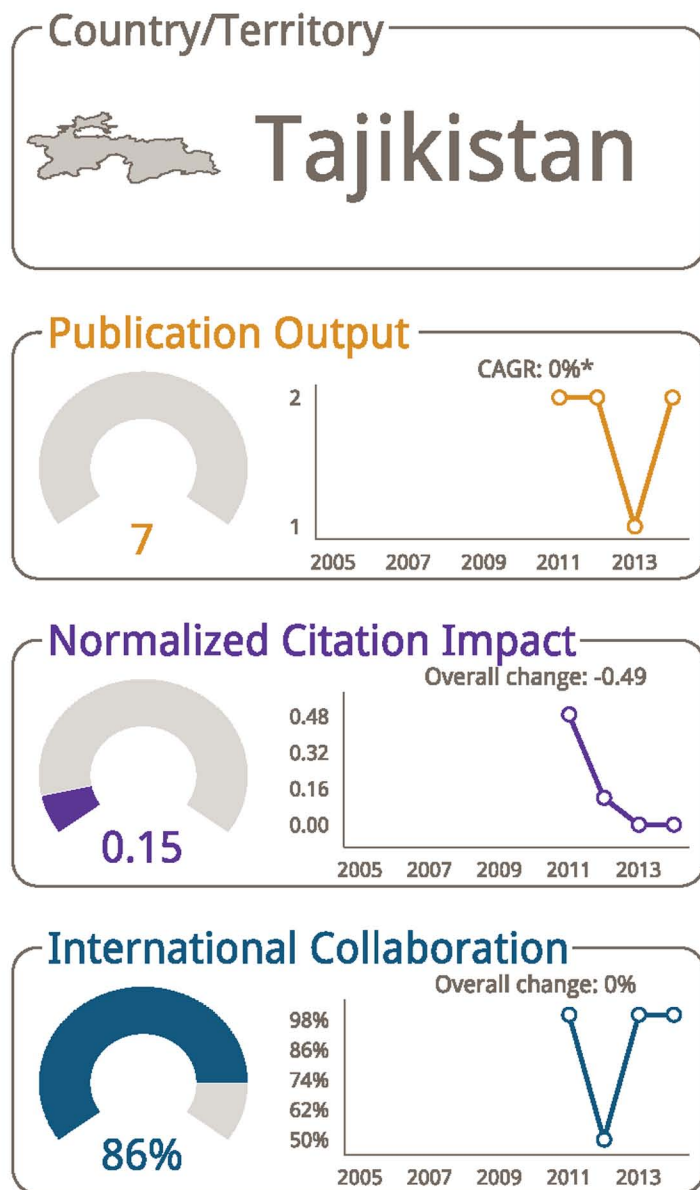


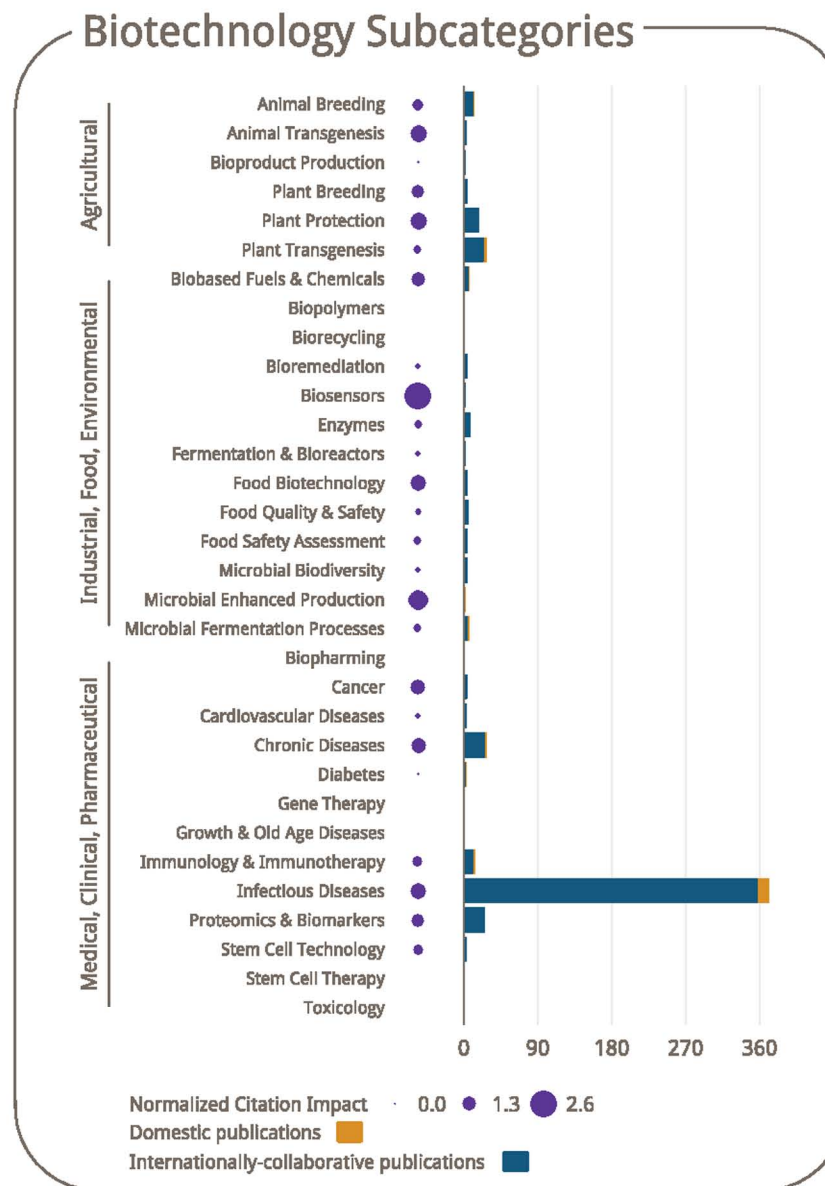
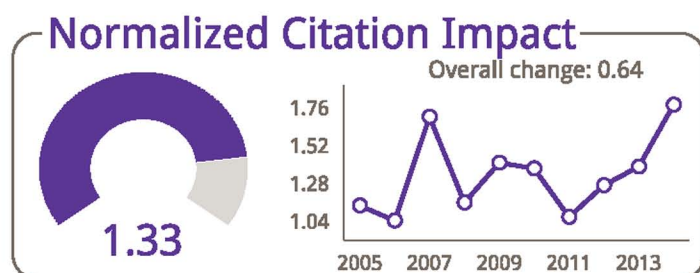
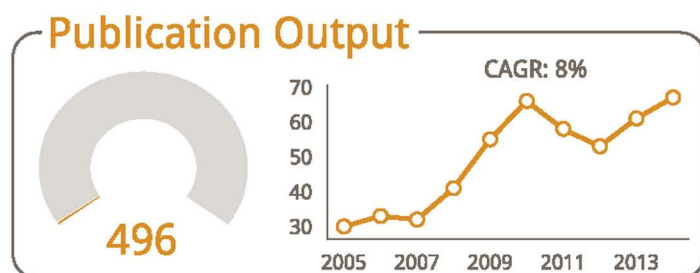


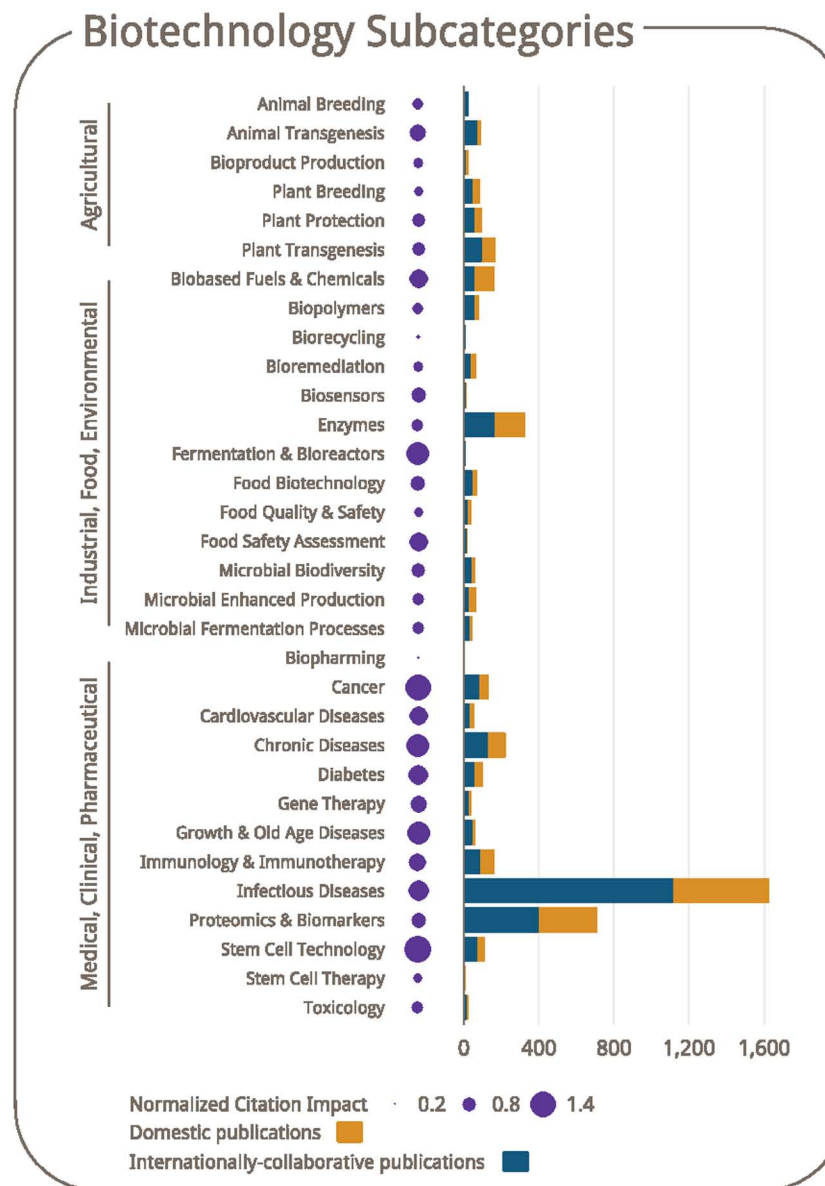
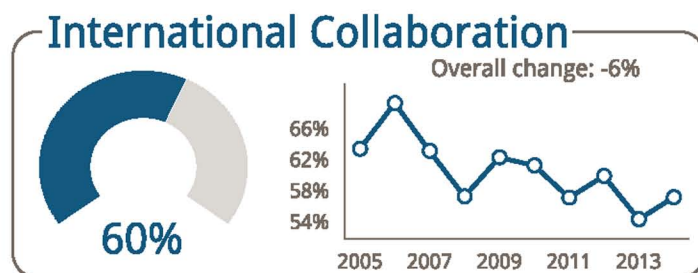
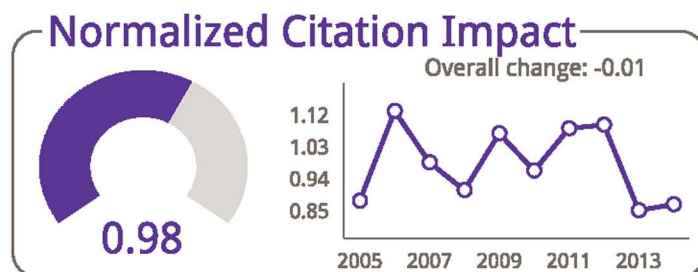
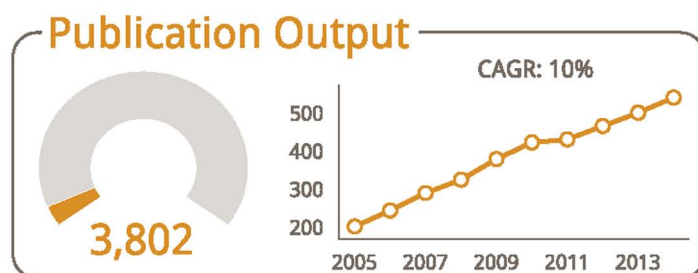


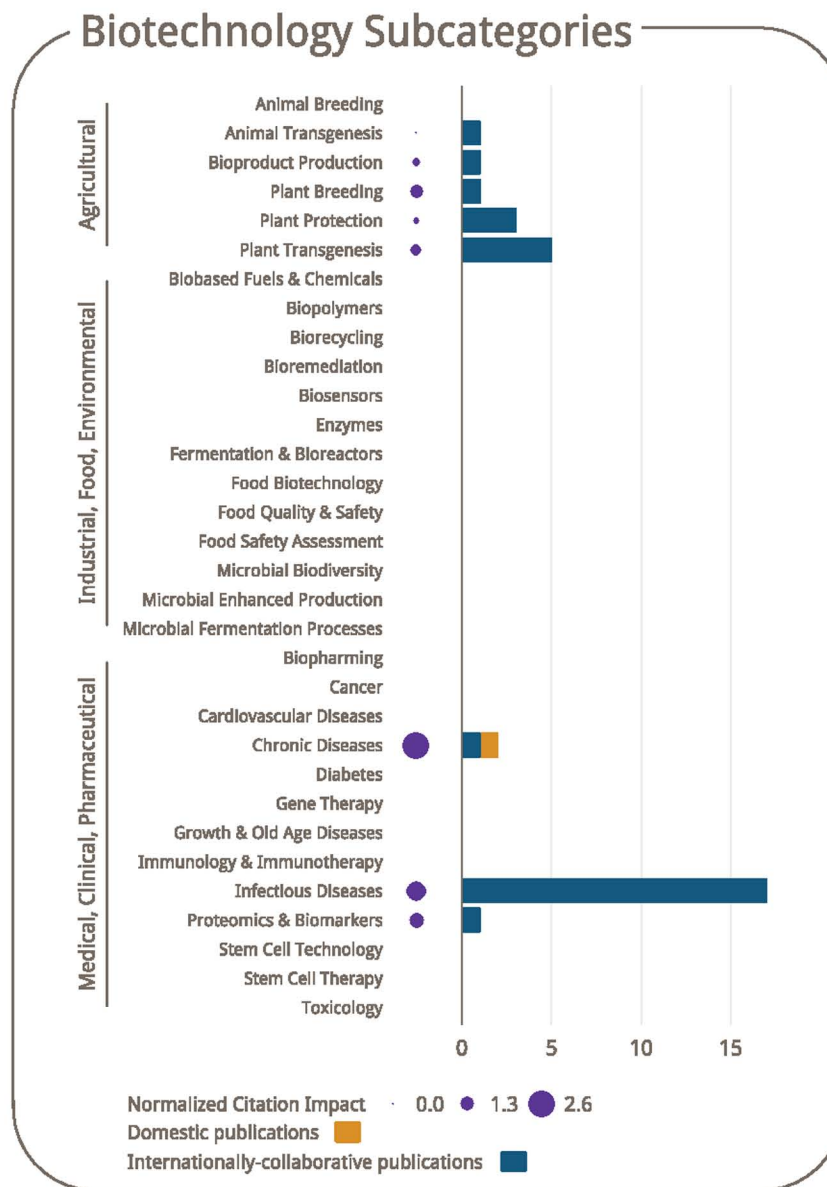
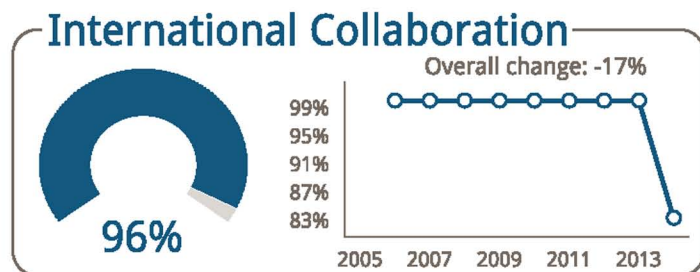
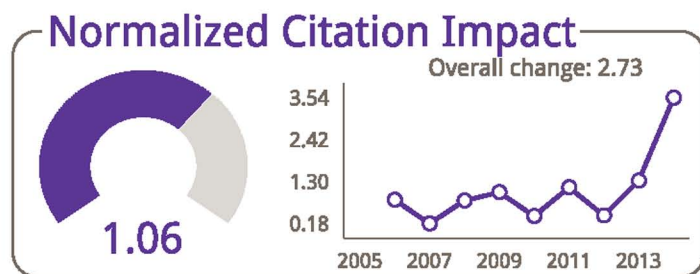
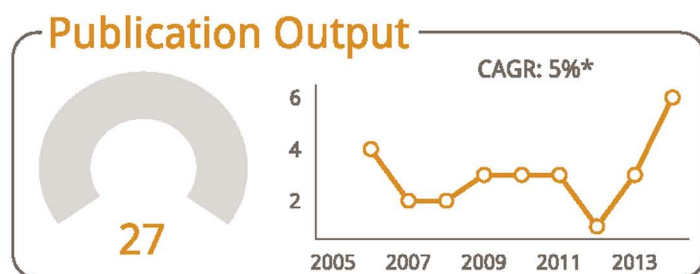












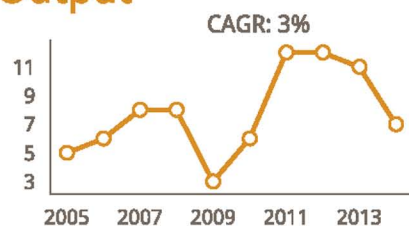


## Country/Territory

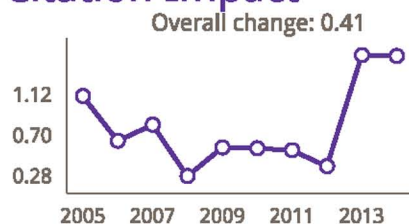


Trinidad and Tobago

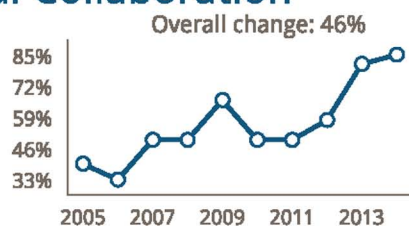
## Publication Output



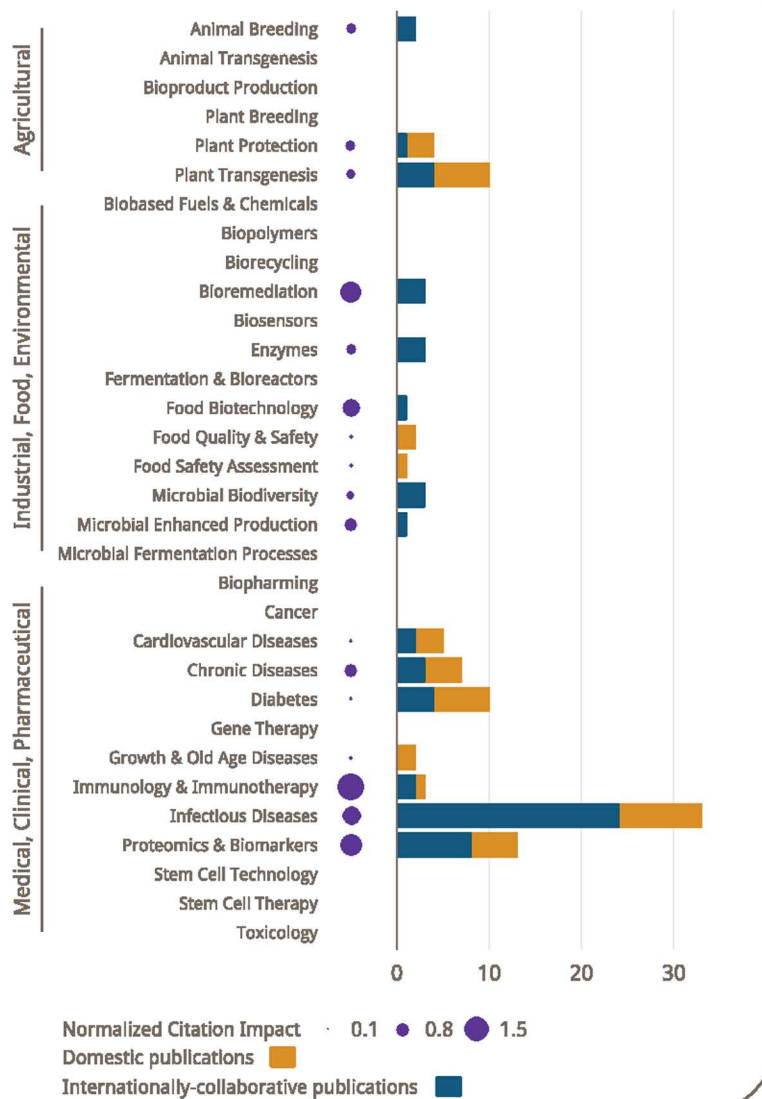
## Normalized Citation Impact



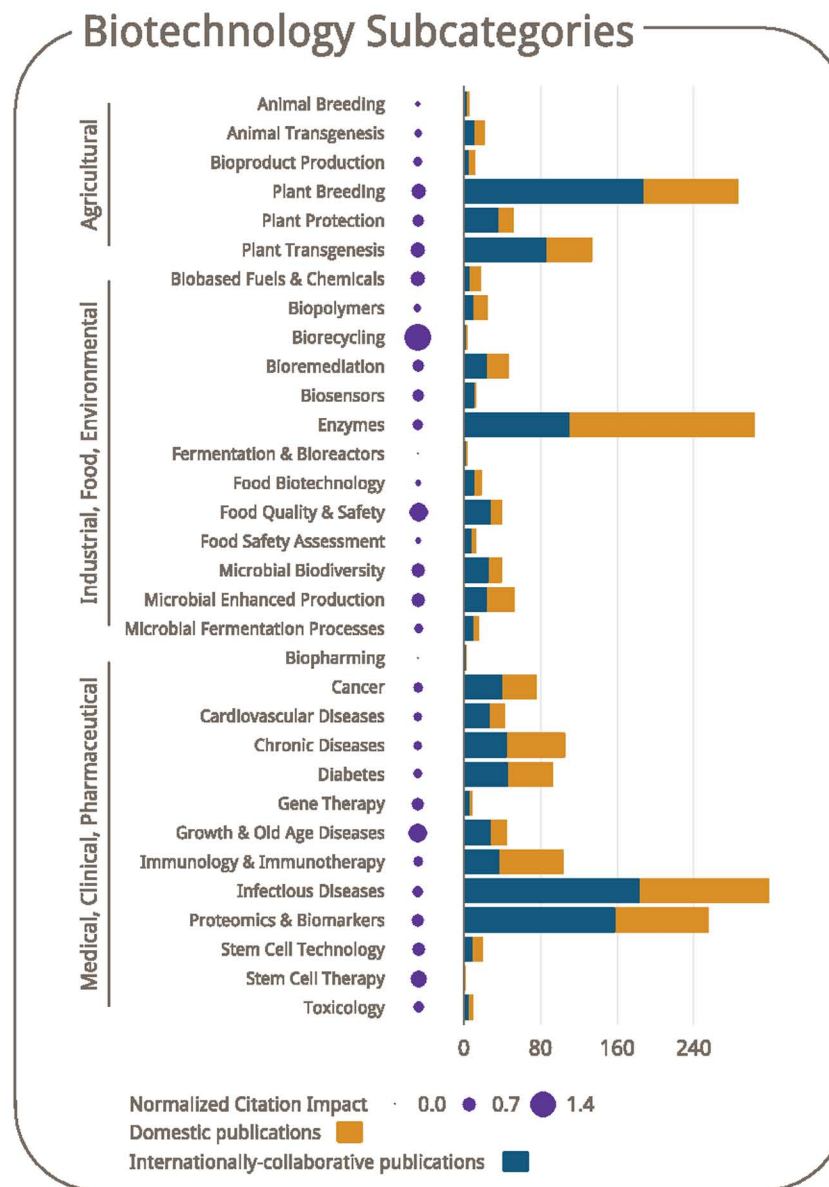
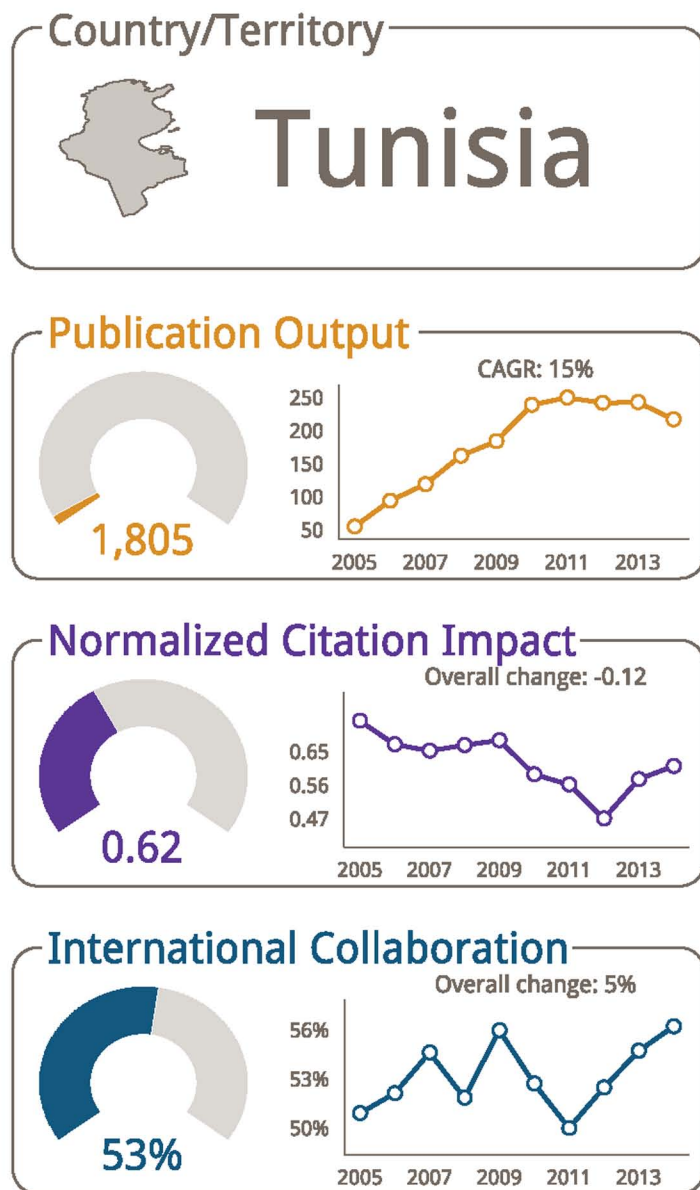
## International Collaboration

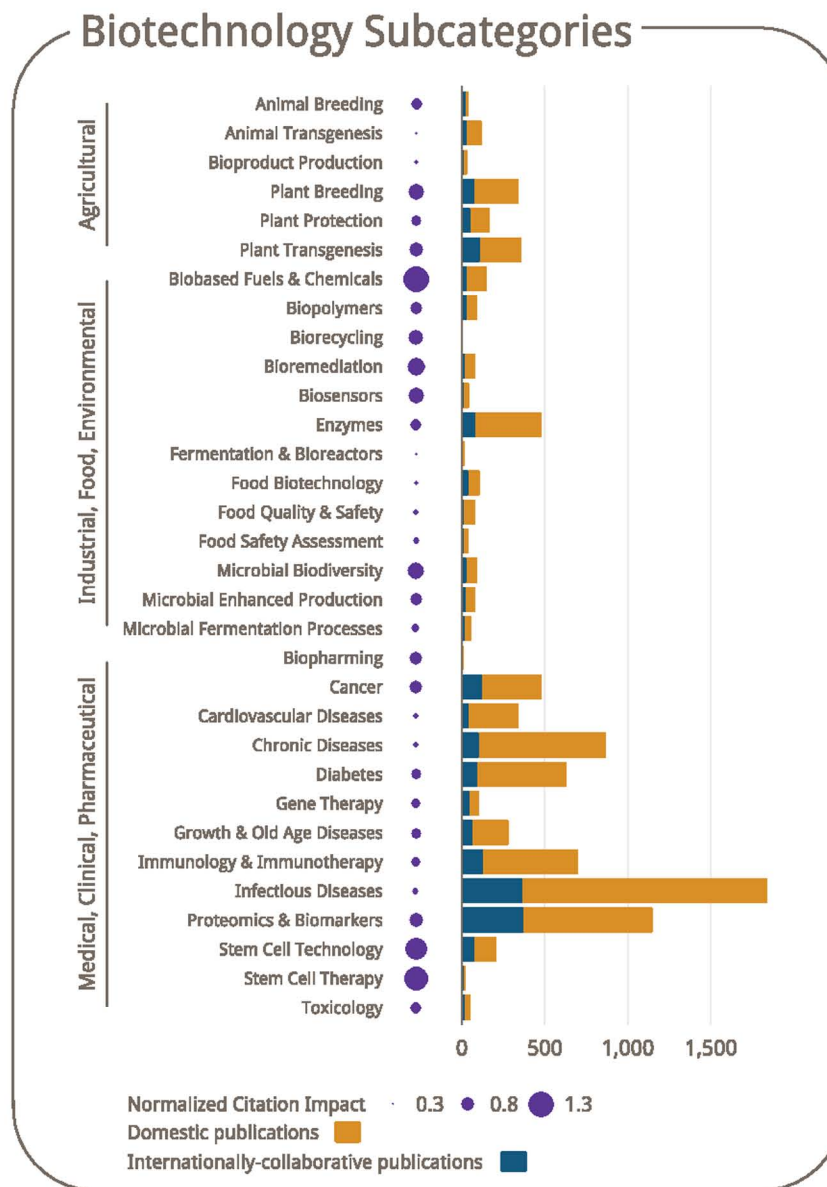
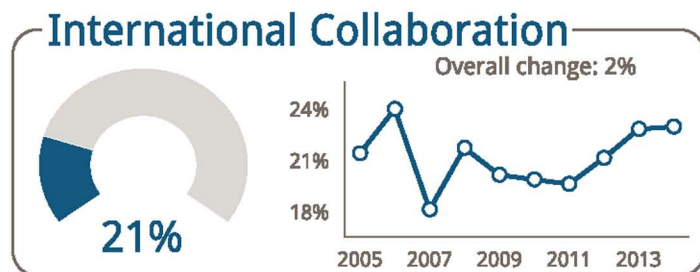
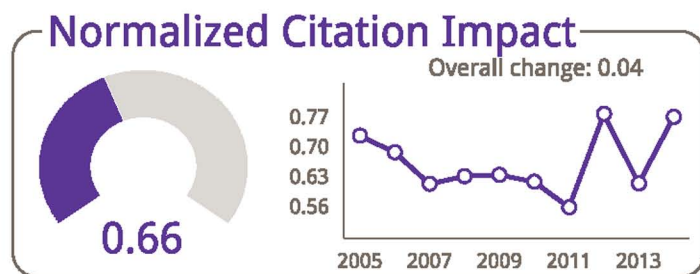
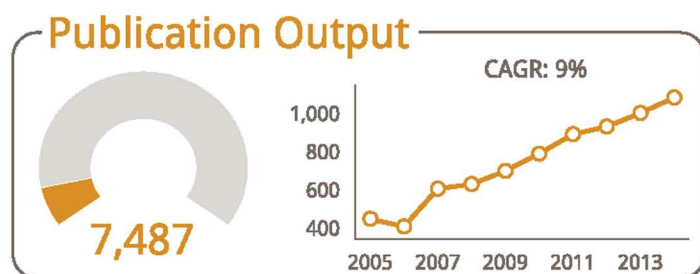


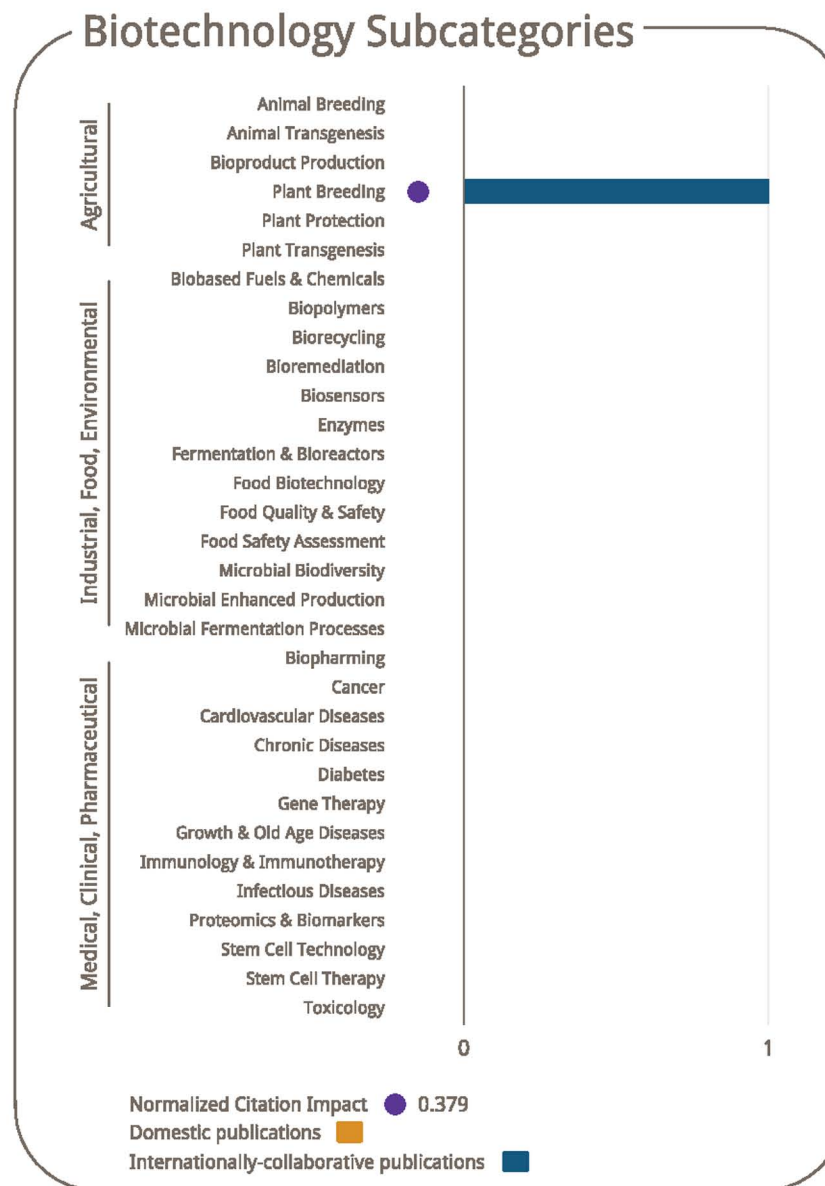
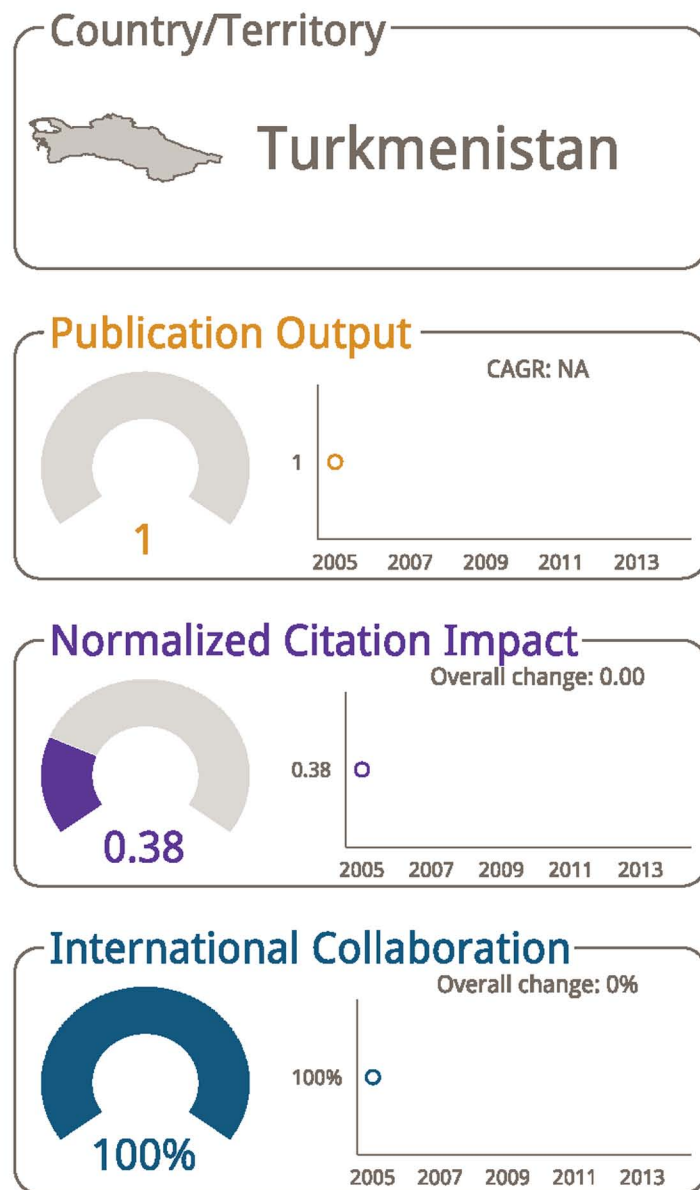
## Biotechnology Subcategories

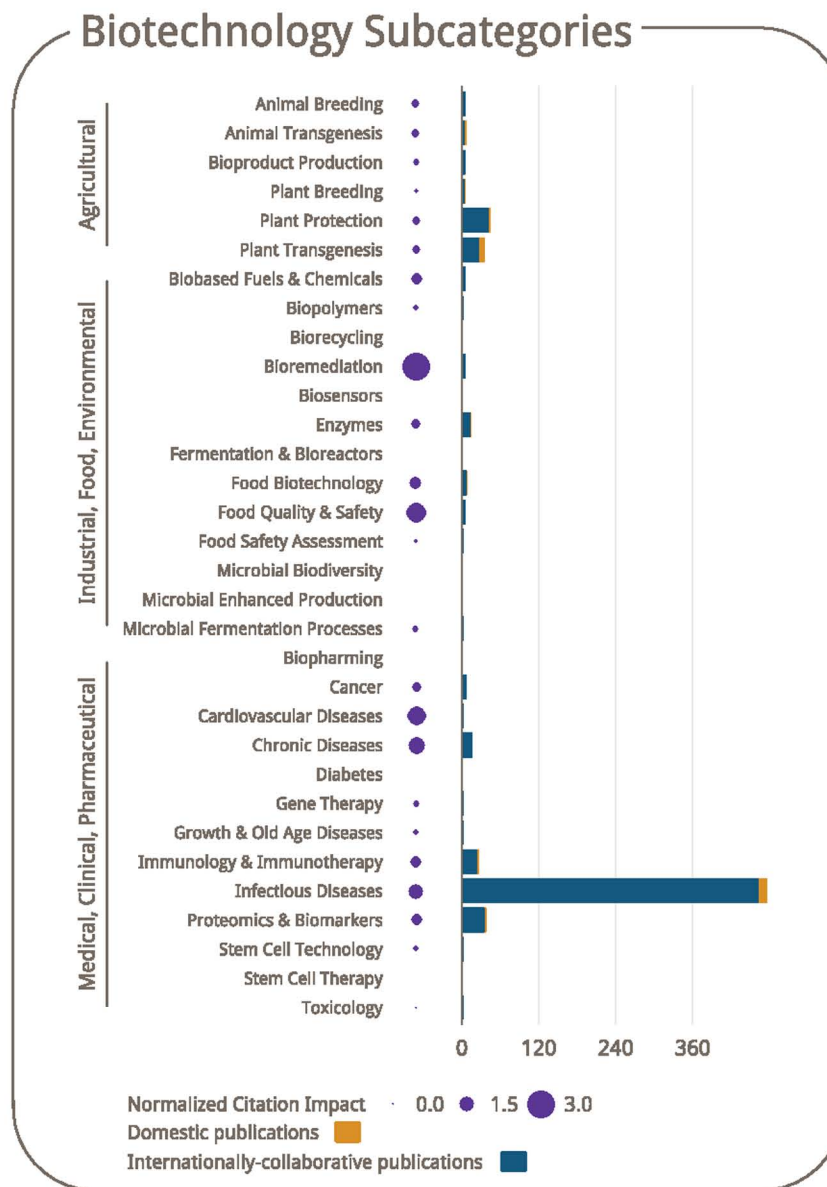
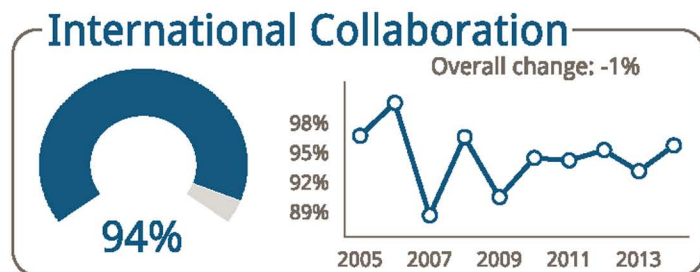
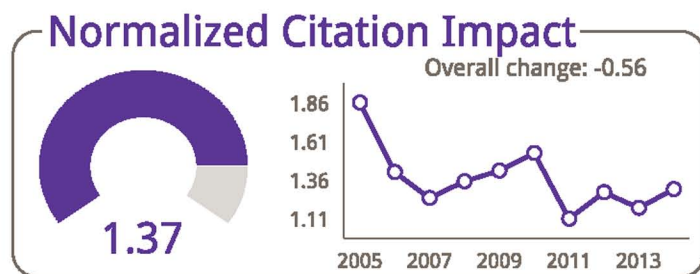
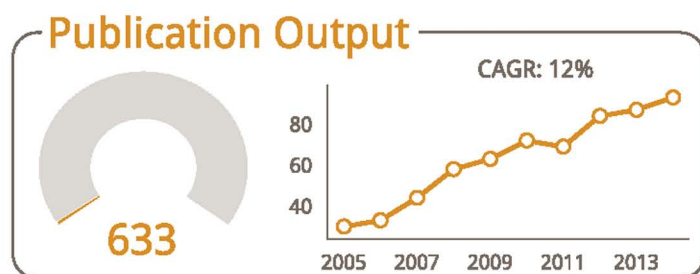










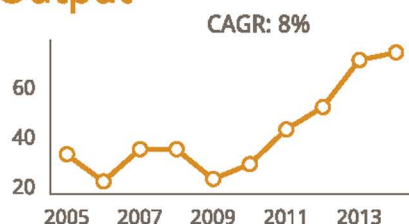


## Country/Territory

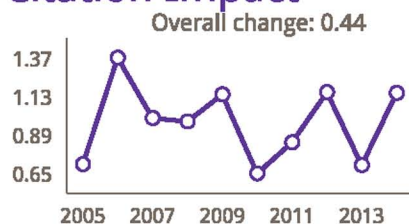


United Arab Emirates

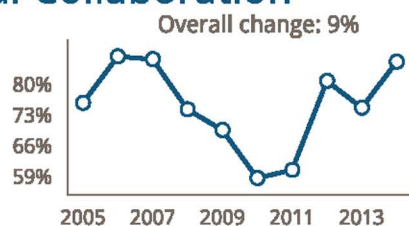
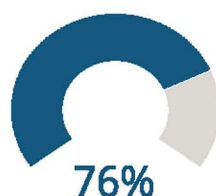
## Publication Output



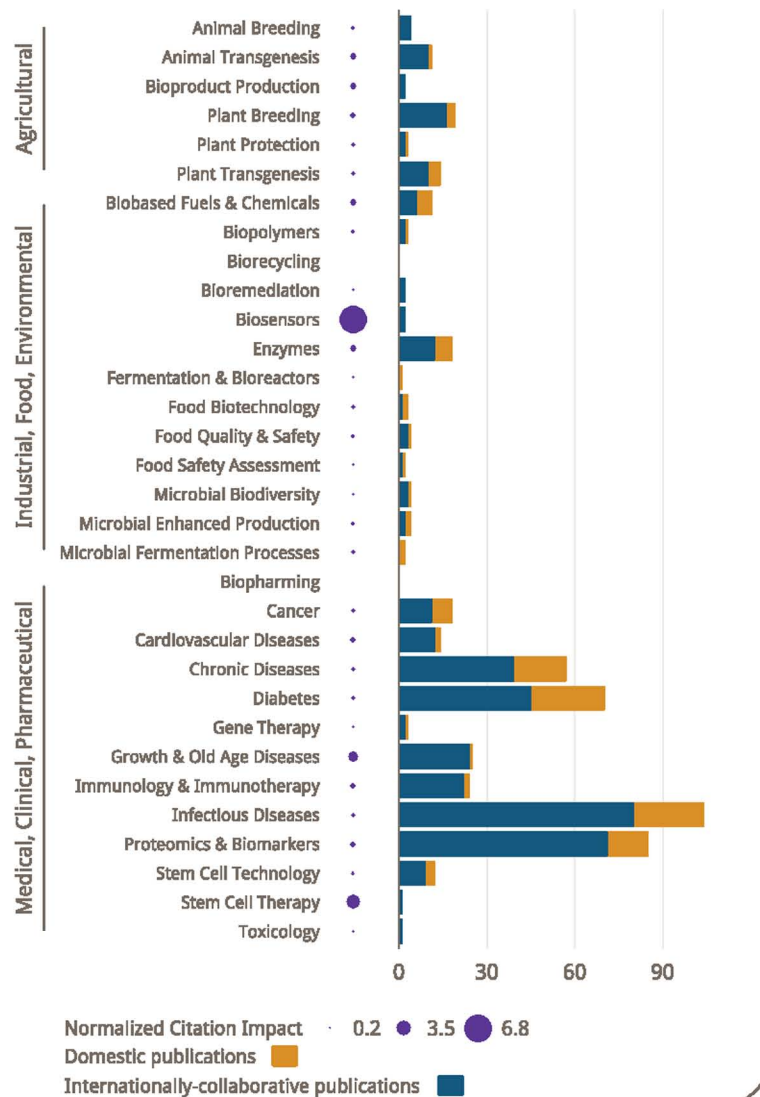
## Normalized Citation Impact



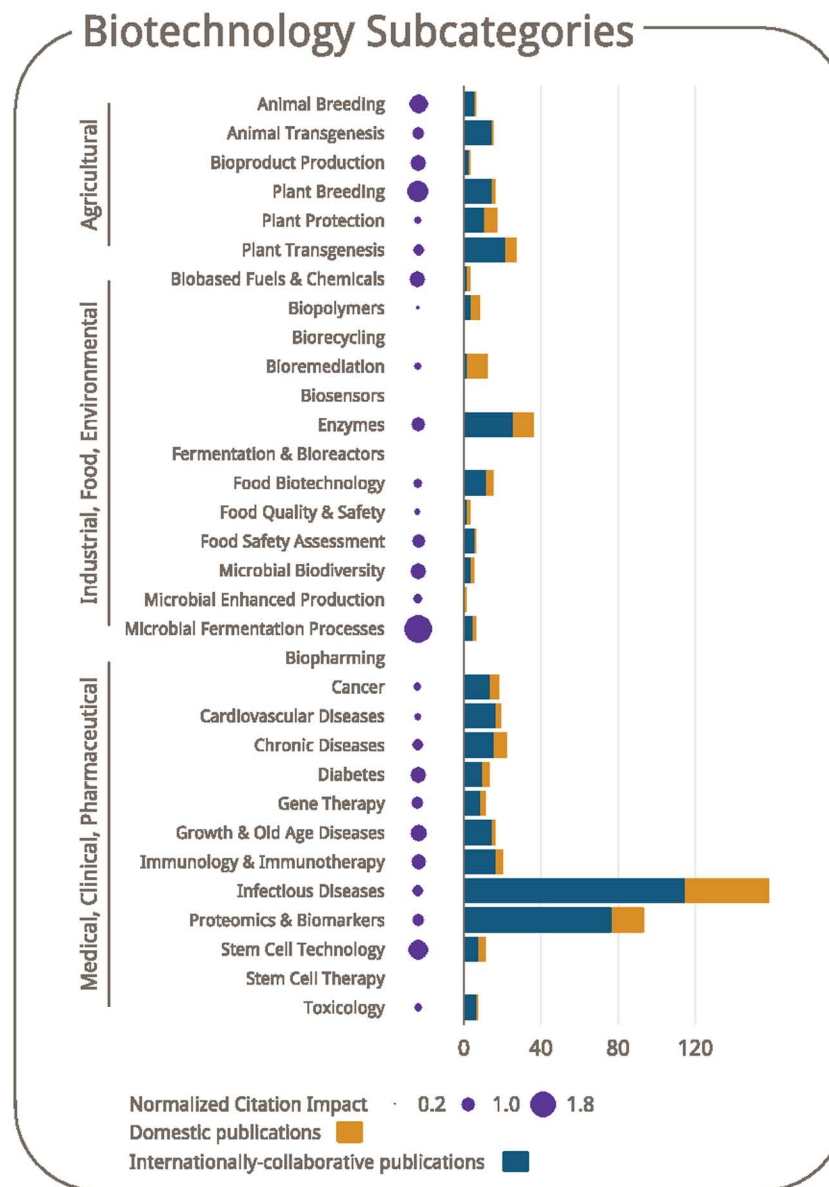
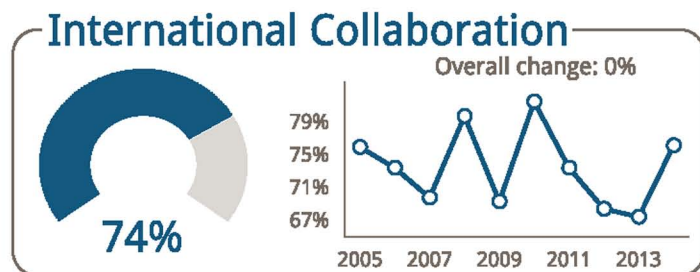
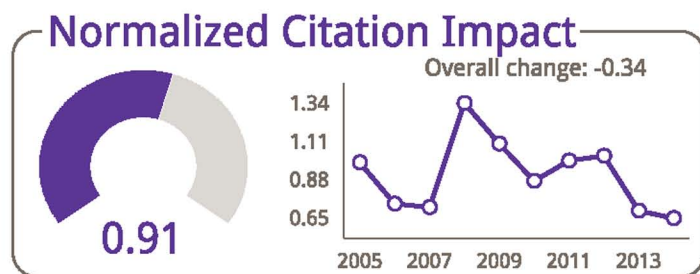
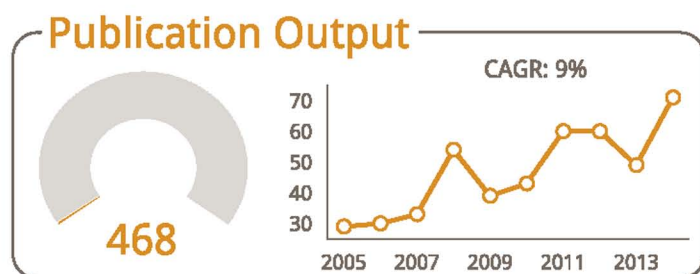
## International Collaboration



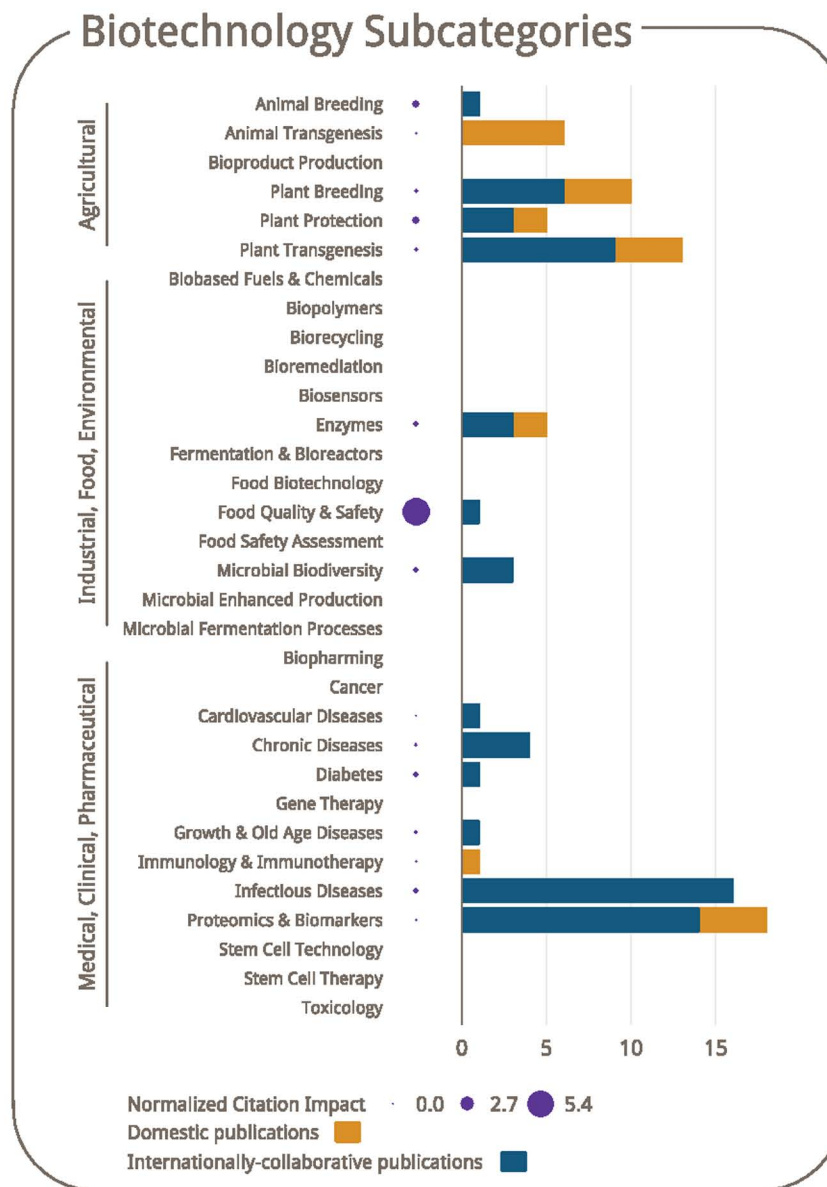
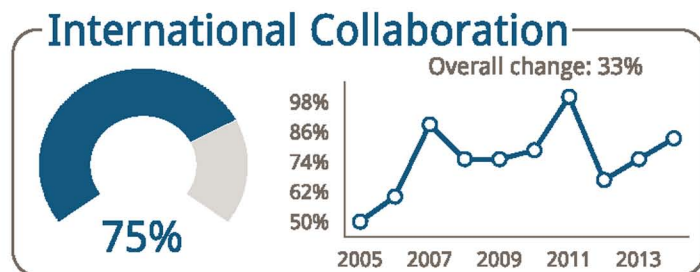
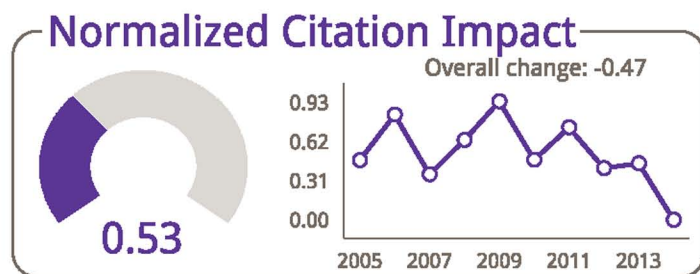
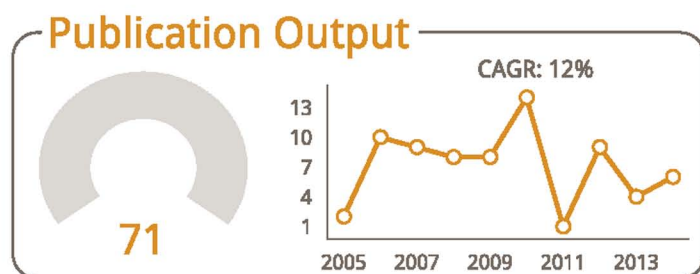
## Biotechnology Subcategories

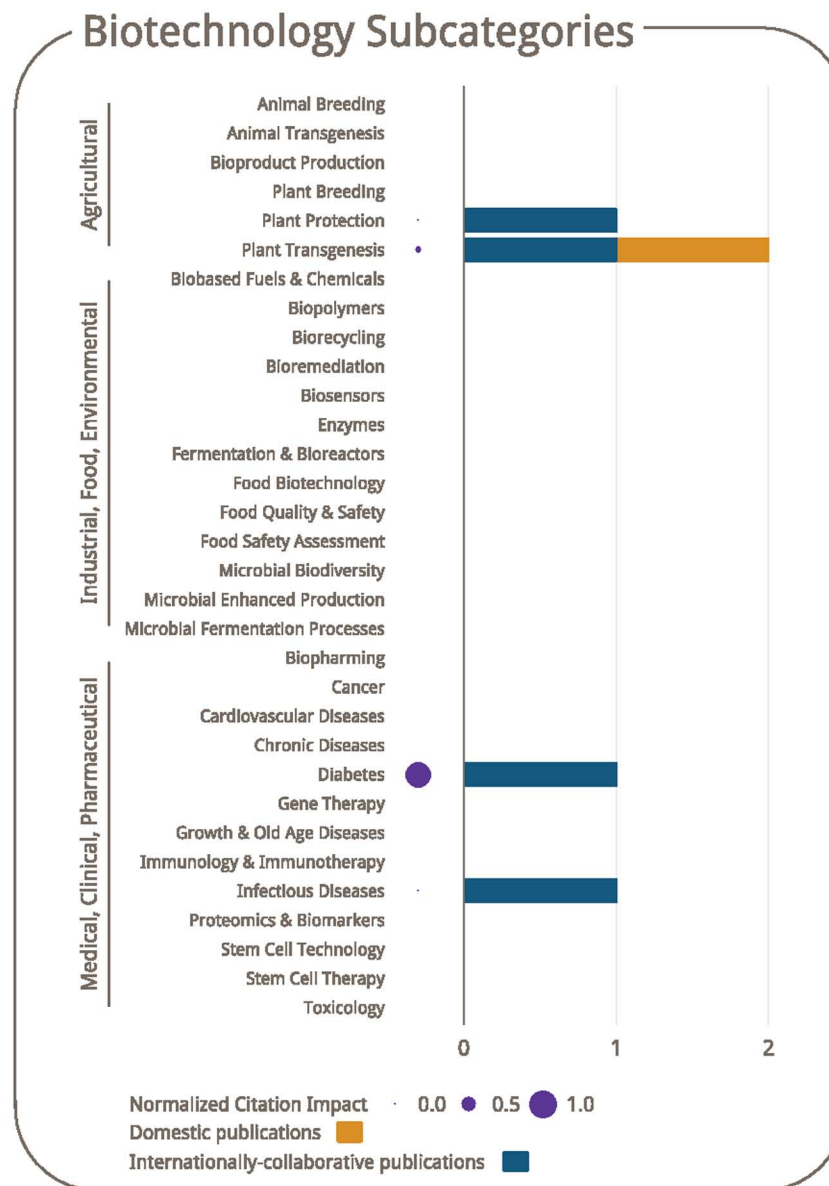
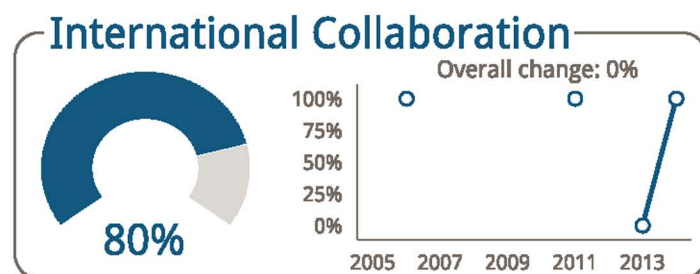
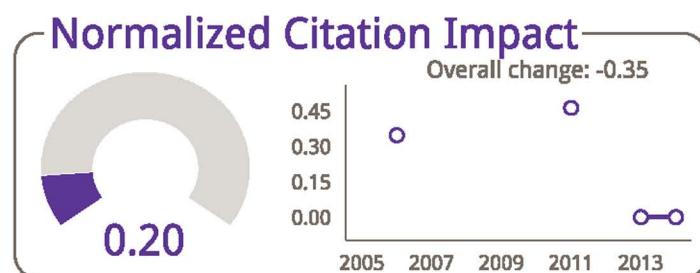
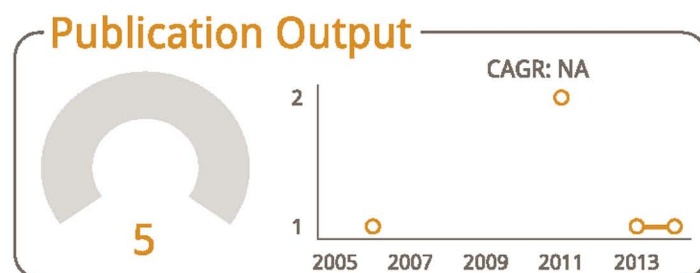


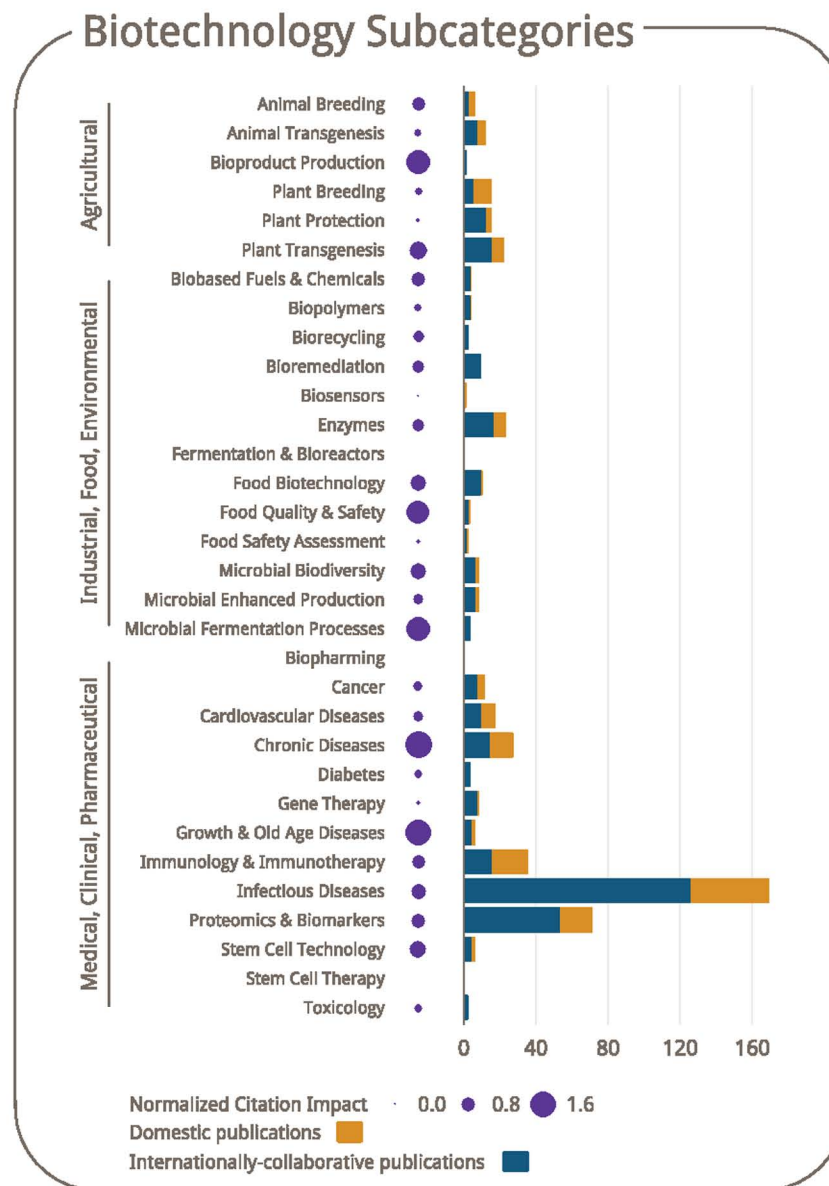
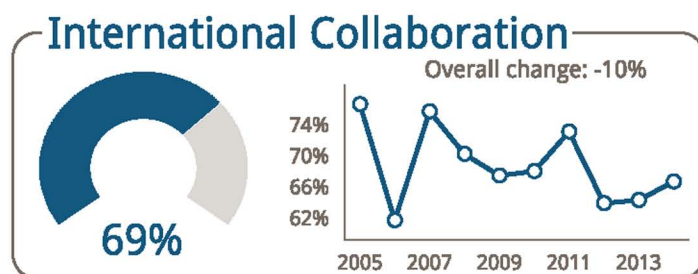
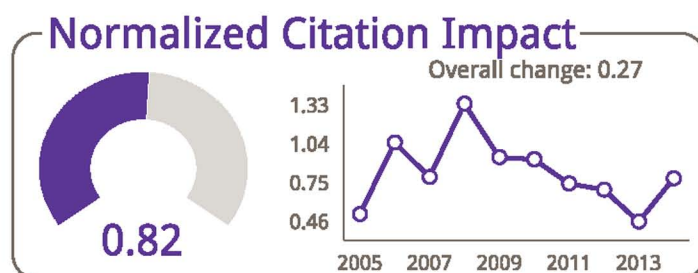
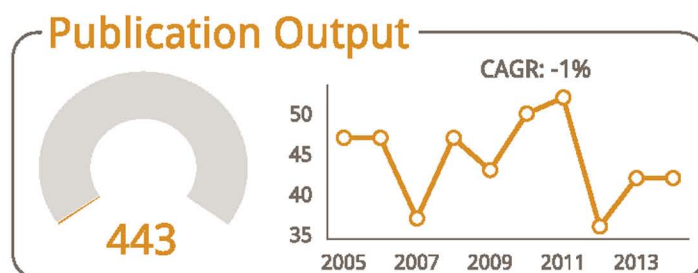


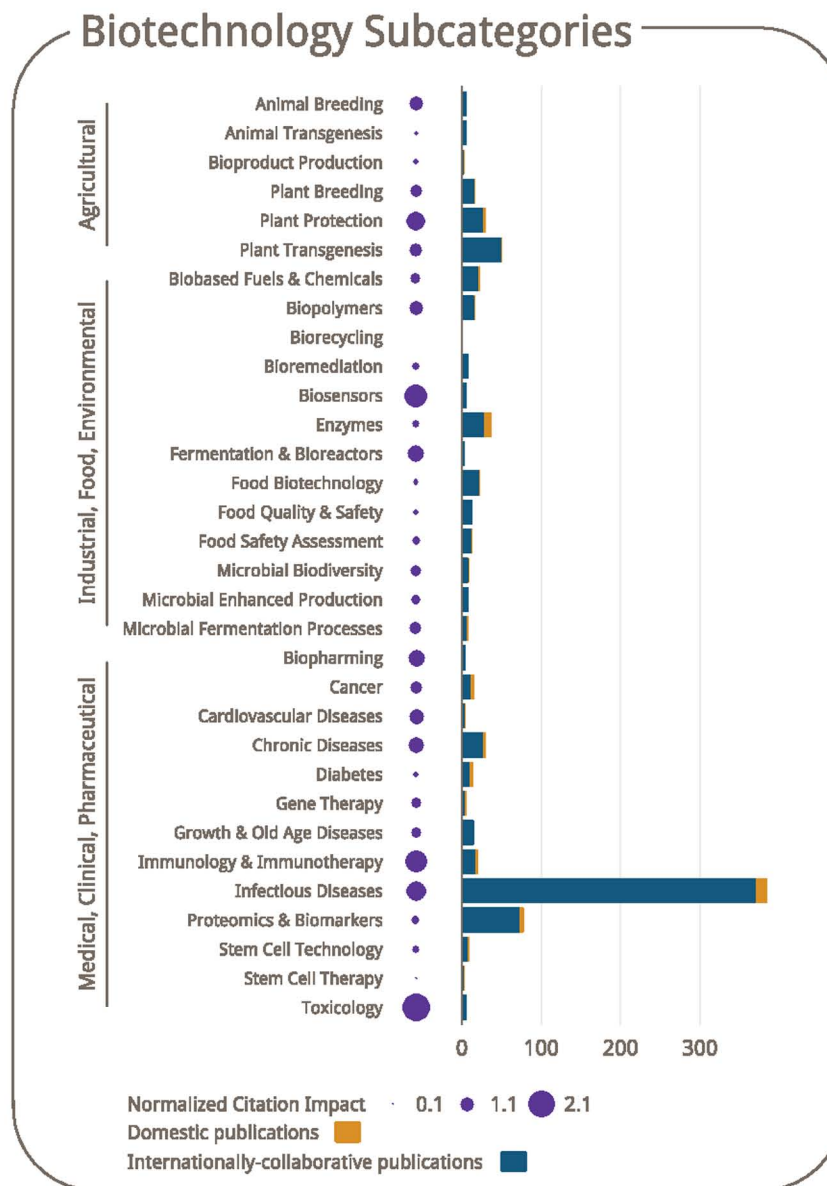
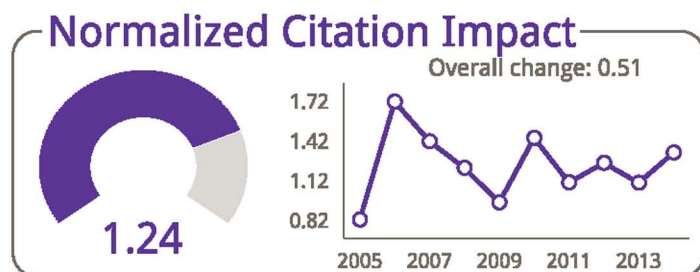
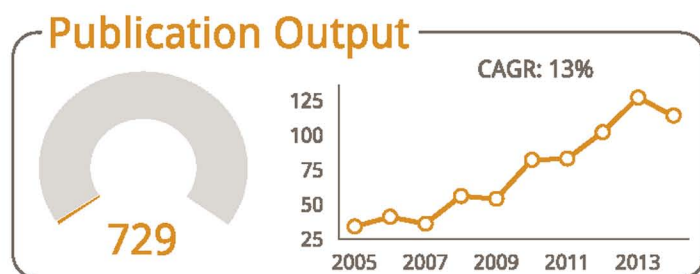


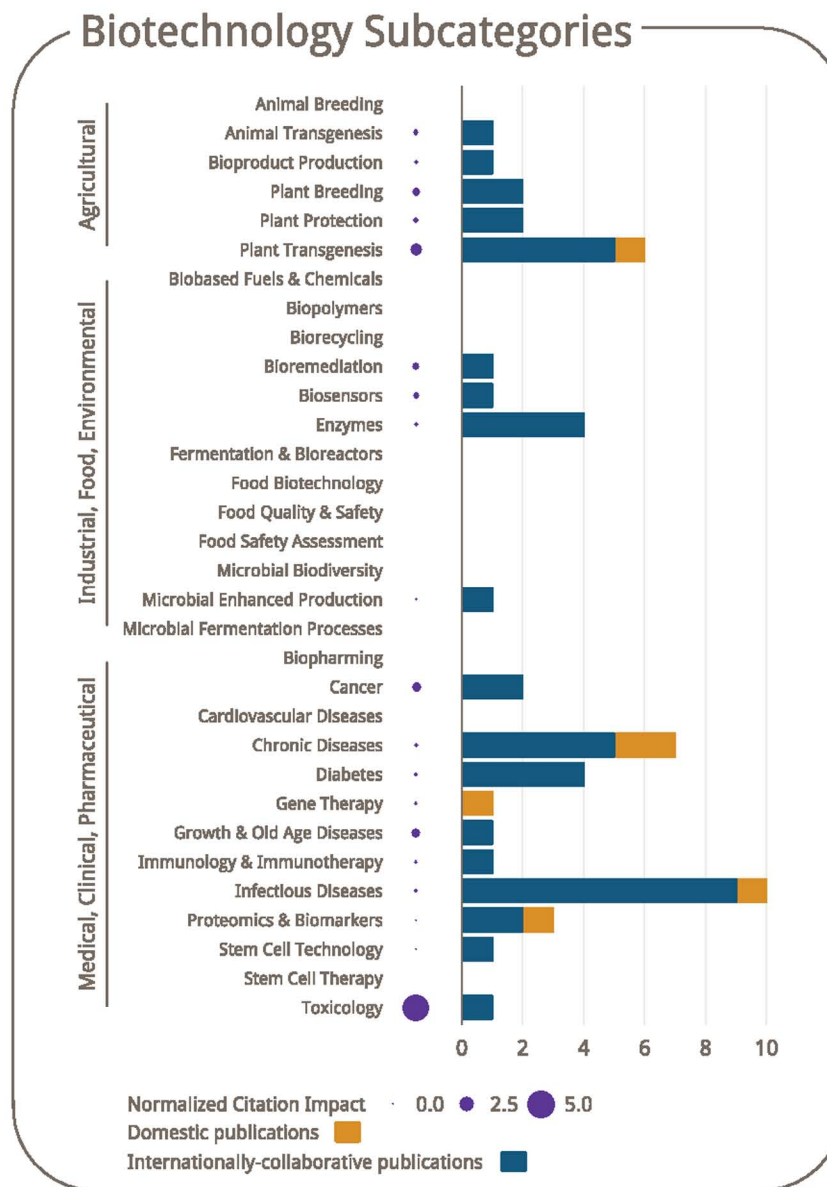
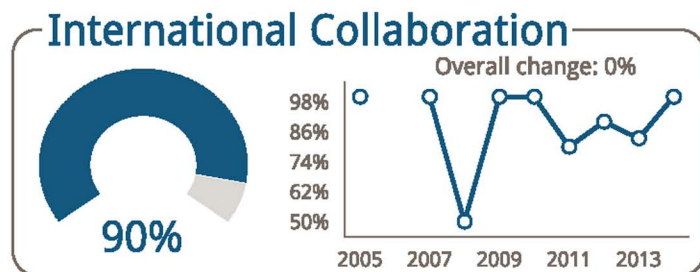
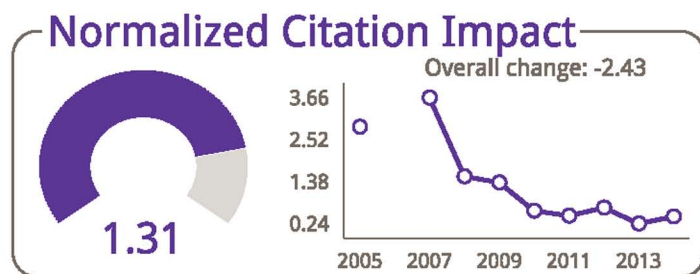
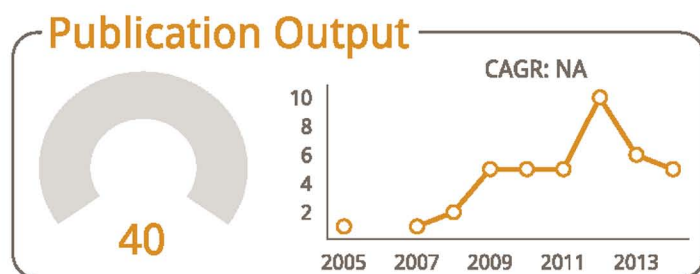




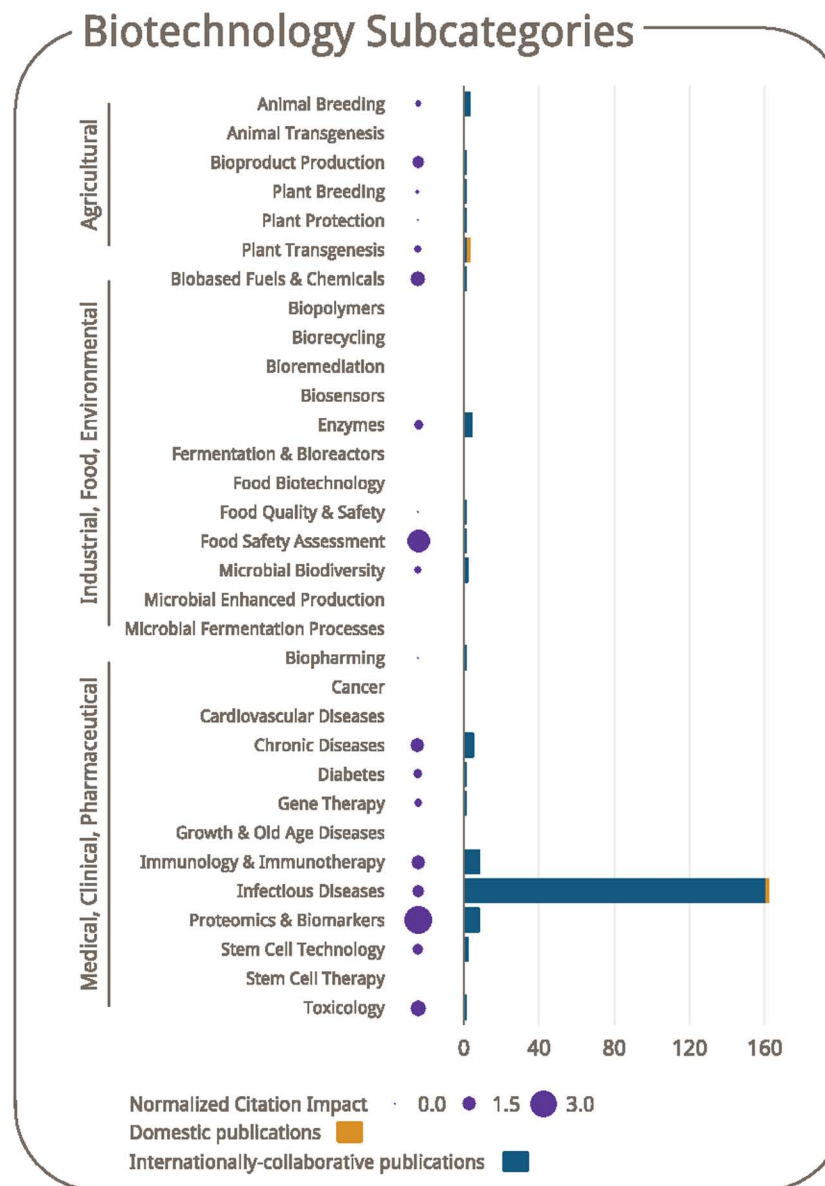
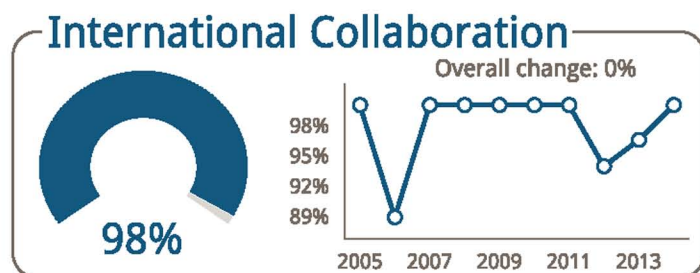
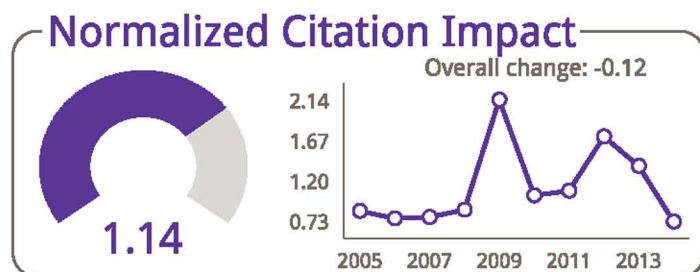
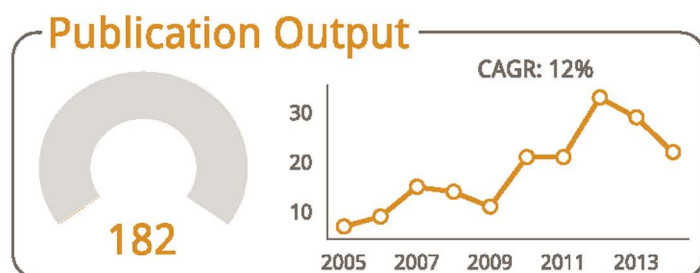




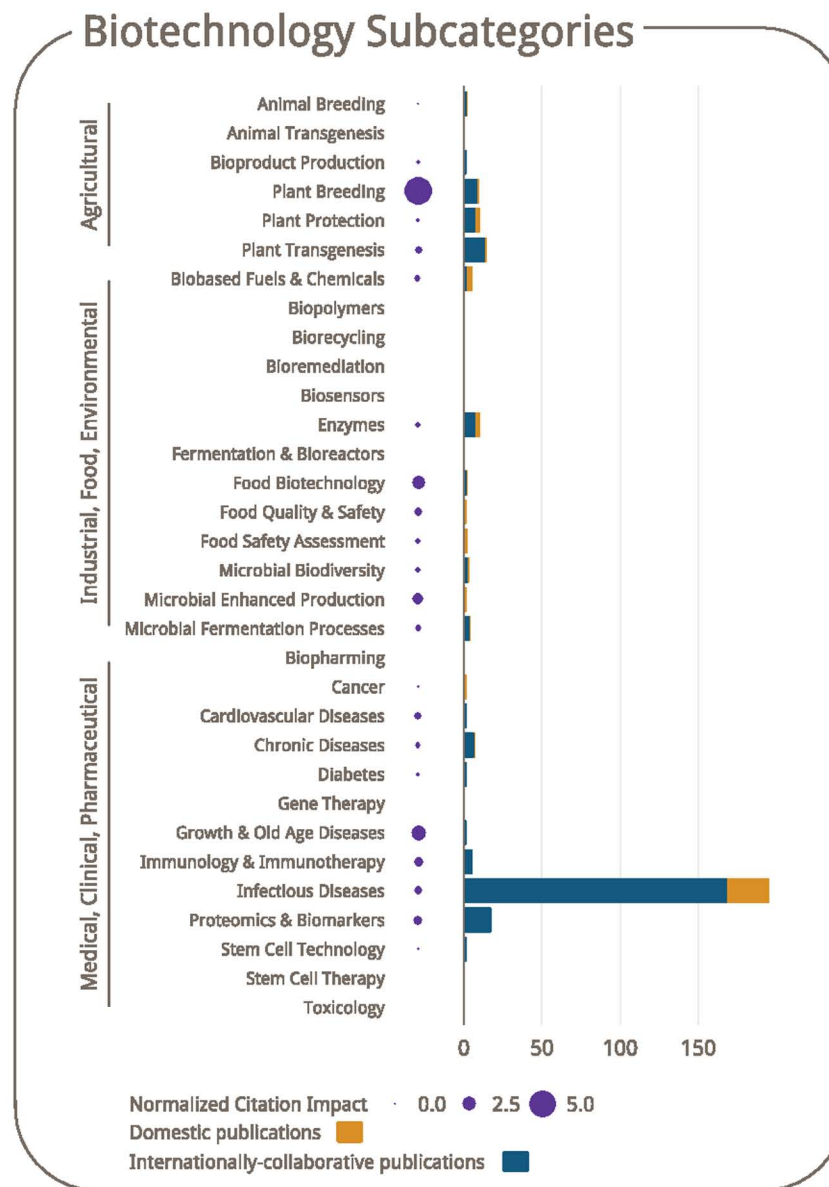
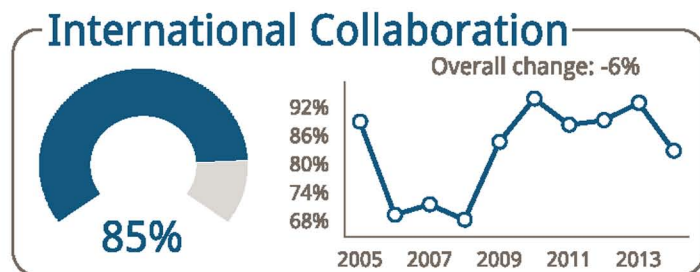
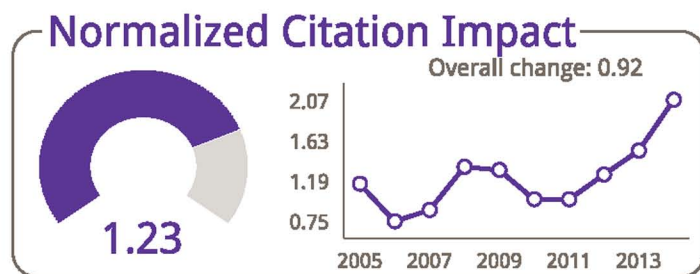
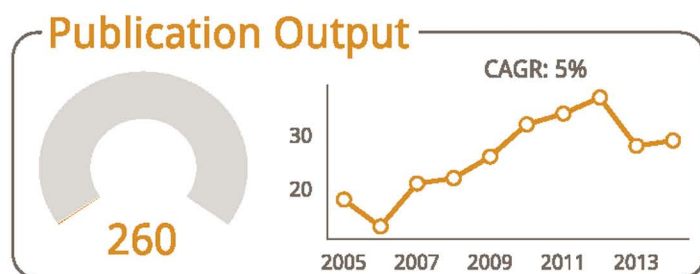








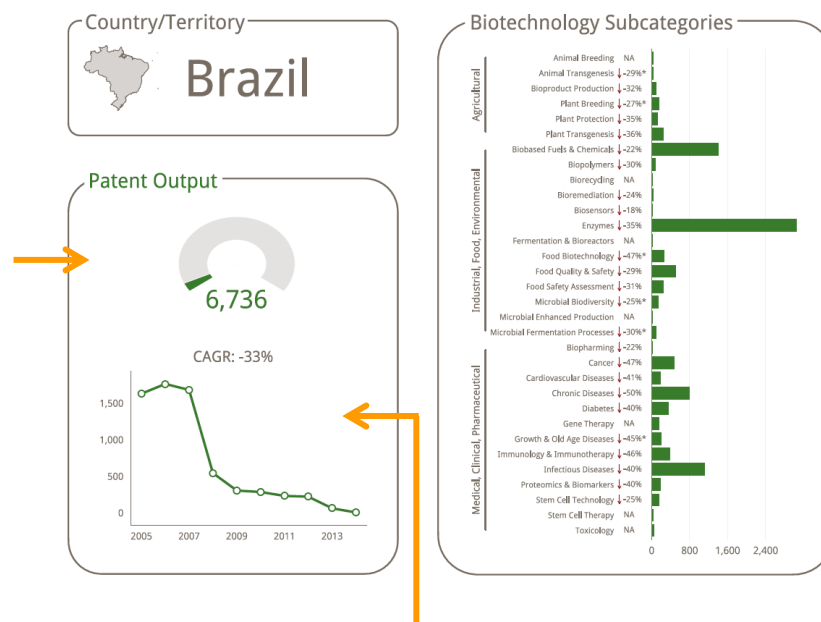




## 7.2 Patent Profiles

The patents profiles are presented for each of the 30 countries in the following layout.

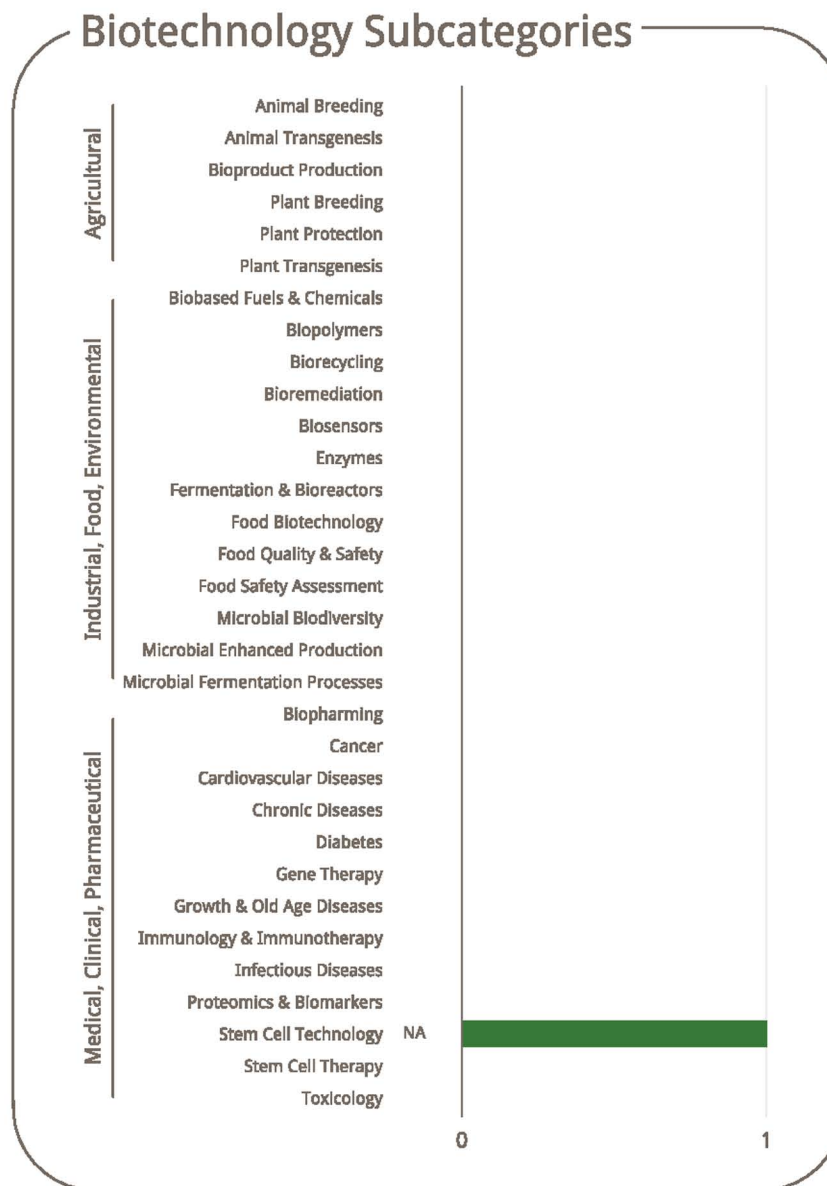
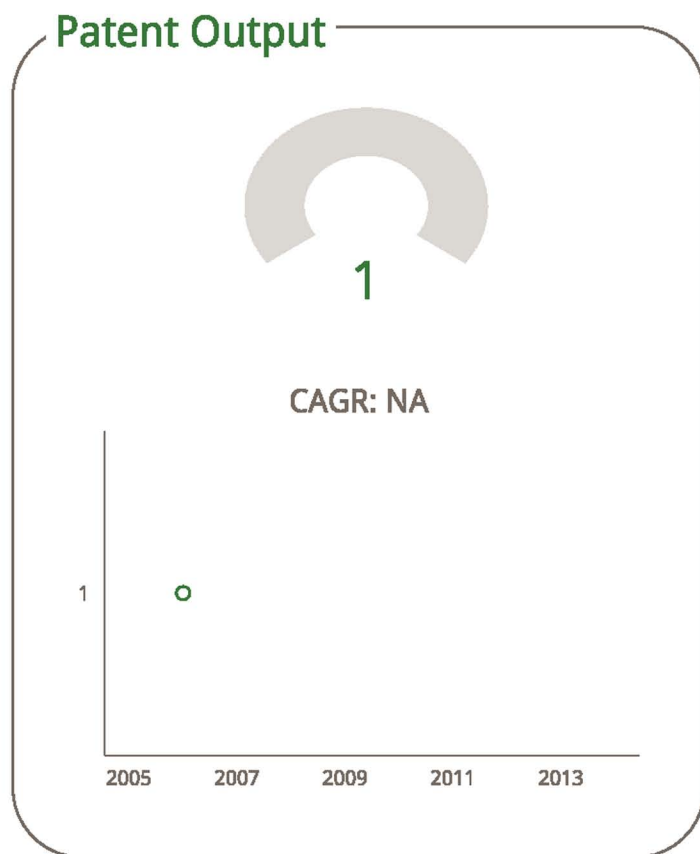
Dials show a country's overall level of patent output, relative to the minimum and maximum values across all 30 countries. For example the minimum publication output for all the countries is 1 while the maximum is 149,339 (China). Brazil's publication output of 6,736 is about one twentieth of the way from the minimum value to the maximum value, and hence its dial is about 5% of the way over.



For each of the 32 **Biotechnology subcategories**, the number patent families is shown in the horizontal barchart, and an green upward/red downward arrow appears next to each subcategory indicates the CAGR of patent families in that subcategory over the period studied. The CAGR value will appear either as NA or with an asterisk when the activity is dis-continuous or only happening in only one time interval but shorter than 2005-2013.

A country's trends in patent output, is shown in the linechart. The growth in publication output is measured by the compound annual growth rate (CAGR, calculated by using  $\frac{\text{End year value}}{\text{Start year value}}^{\frac{1}{\text{\# of years}}} - 1$ ).

The start year was the earliest year of patent filing, and the end year was the most recent patent filing year up to 2013. The CAGR was not calculated (shown as "NA") for those countries which had sporadic (dis-continuous) patent filings. The asterisked CAGRs refer to the cases when the continuous activity was observed in only one time interval but shorter than 2005-2013.



## Country/Territory



# Argentina

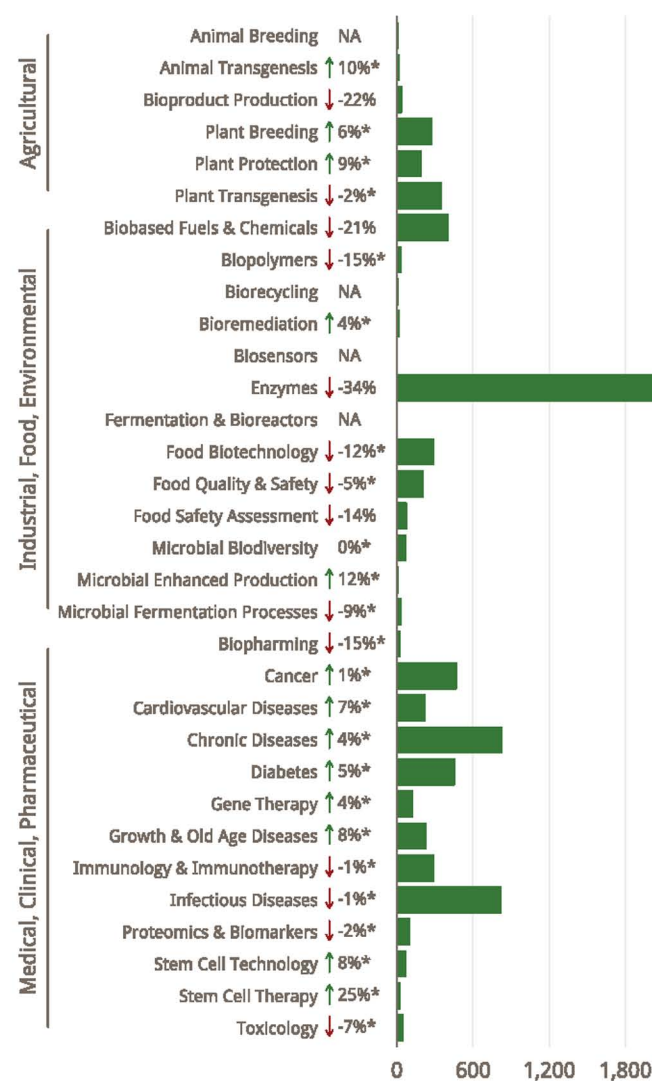
## Patent Output



CAGR: -34%



## Biotechnology Subcategories



## Country/Territory



# Brazil

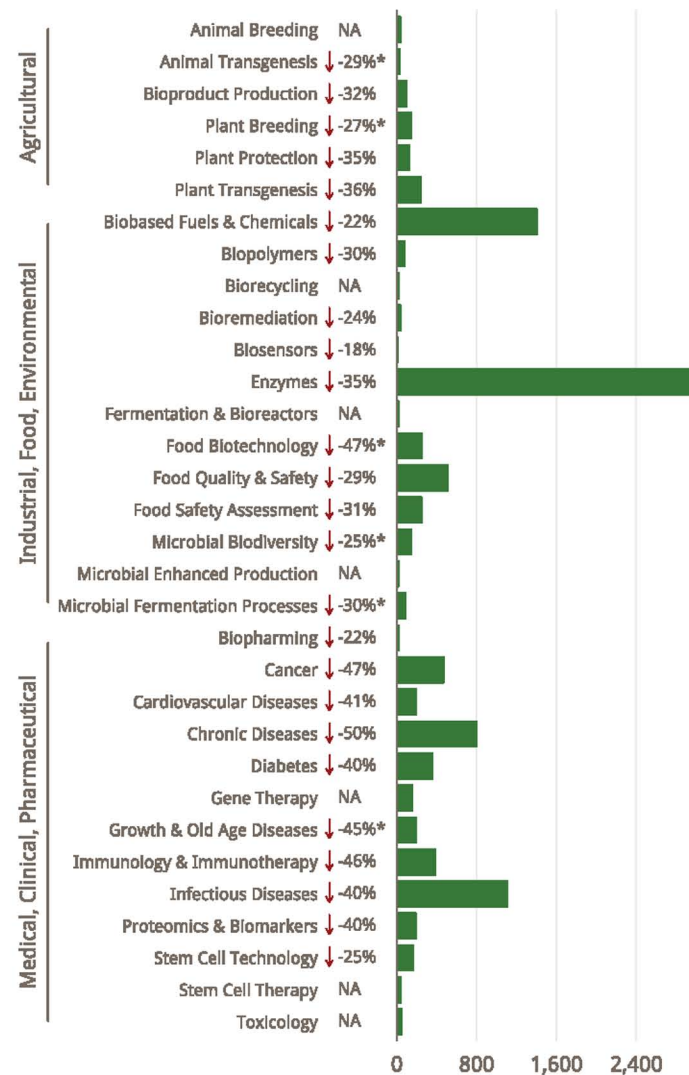
## Patent Output



CAGR: -33%



## Biotechnology Subcategories



## Country/Territory

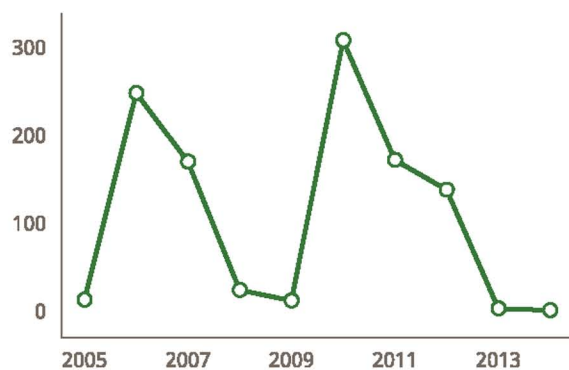


# Chile

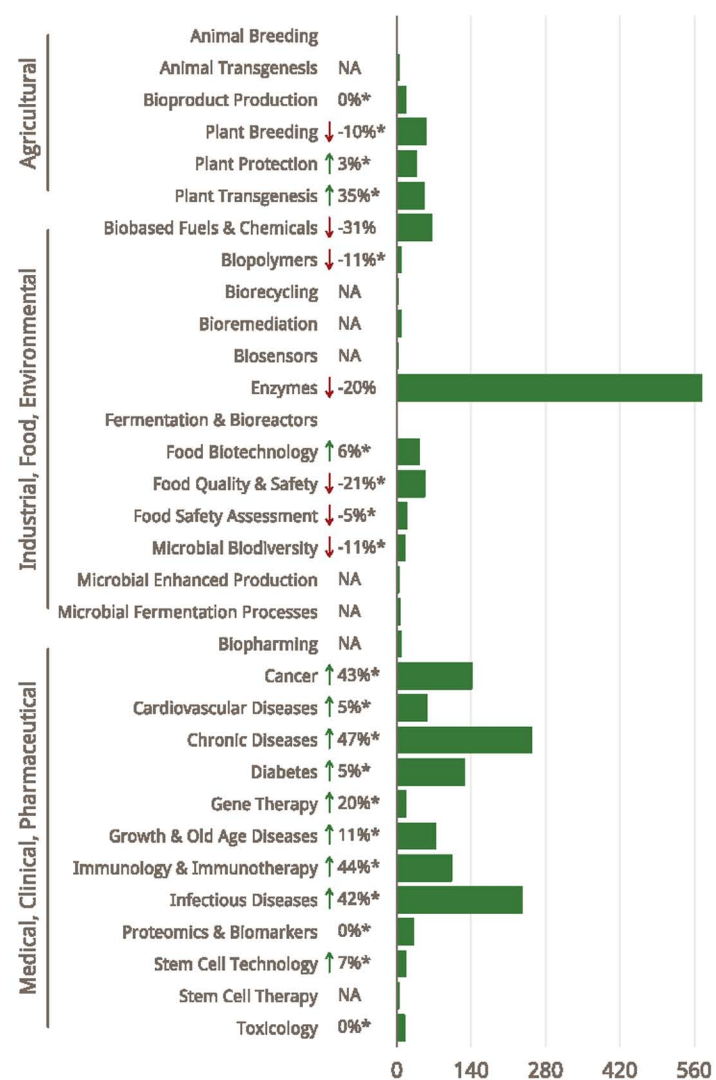
## Patent Output



CAGR: -17%



## Biotechnology Subcategories





## Country/Territory



# China

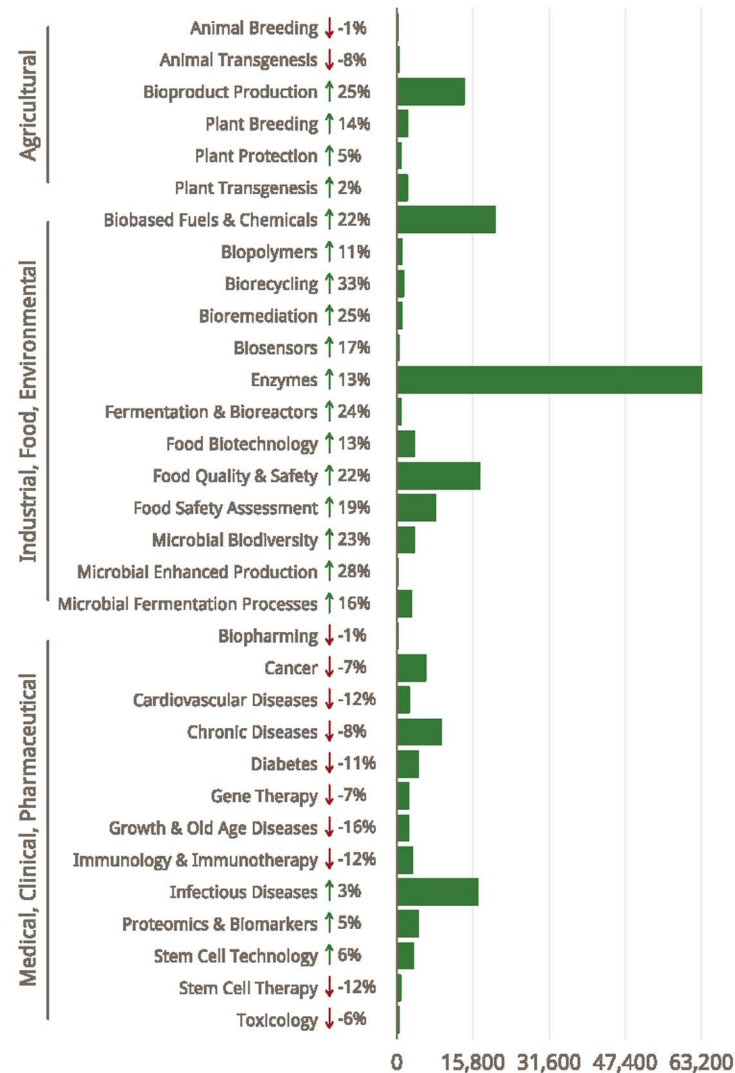
## Patent Output



CAGR: 15%



## Biotechnology Subcategories



## Country/Territory

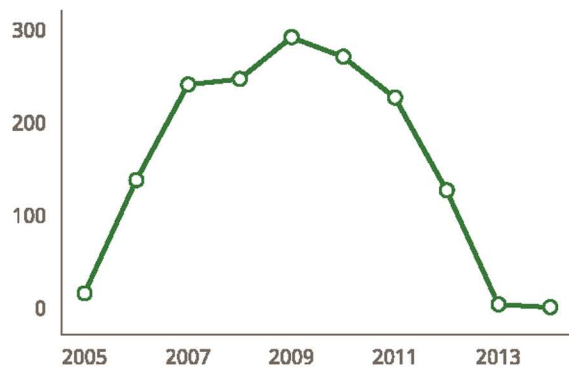


# Colombia

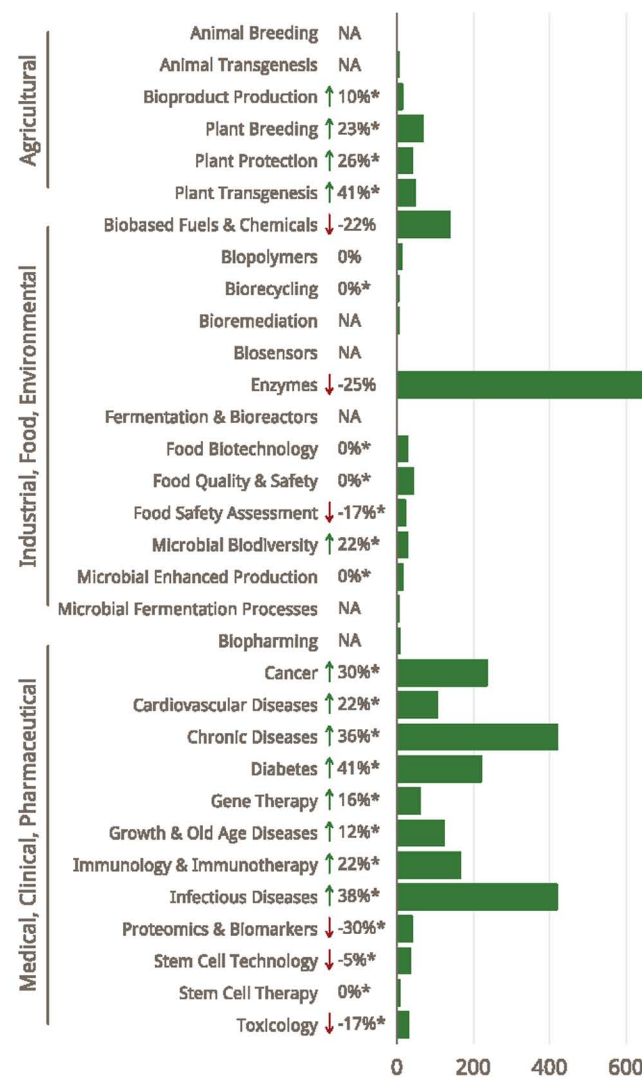
## Patent Output



CAGR: -16%



## Biotechnology Subcategories



## Country/Territory

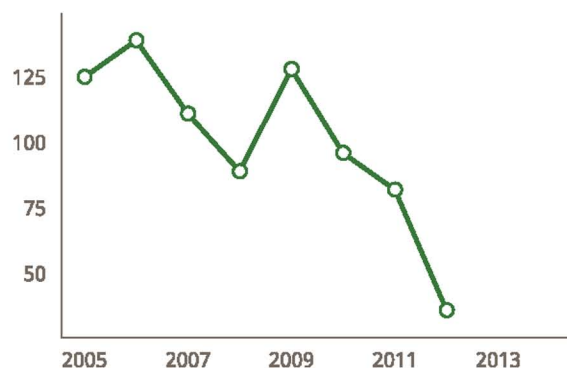


Costa Rica

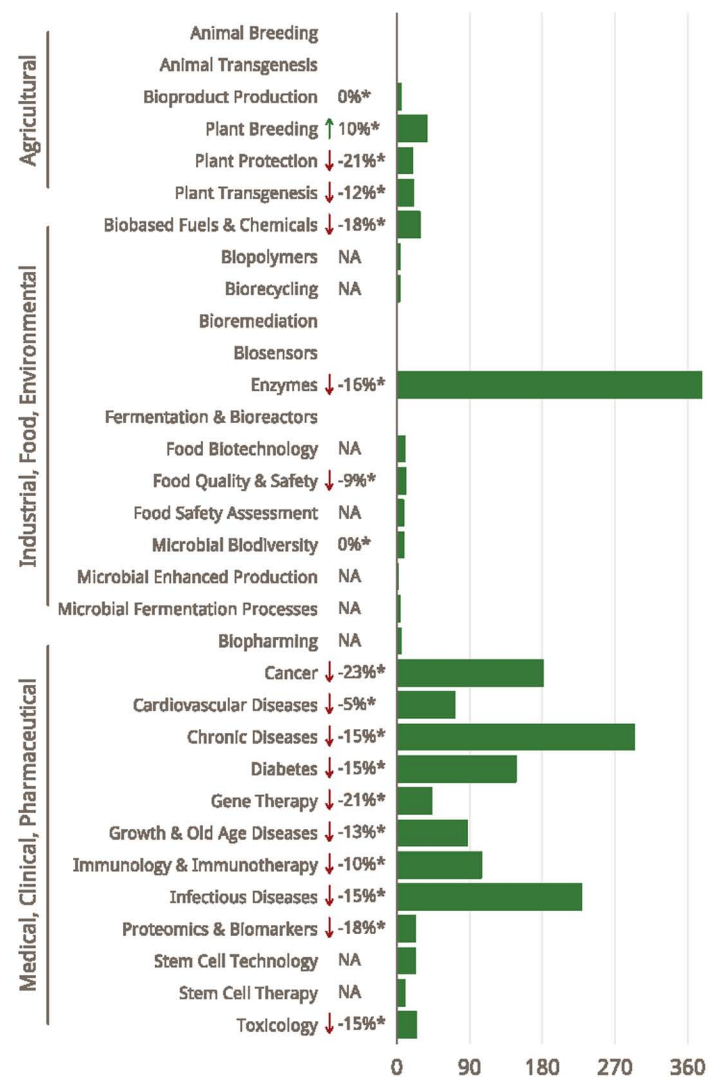
## Patent Output



CAGR: -14%\*



## Biotechnology Subcategories



## Country/Territory

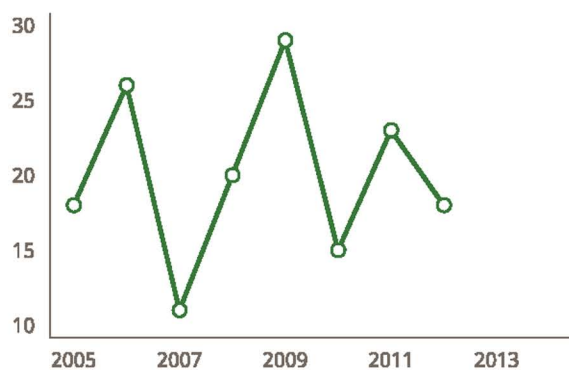


# Cuba

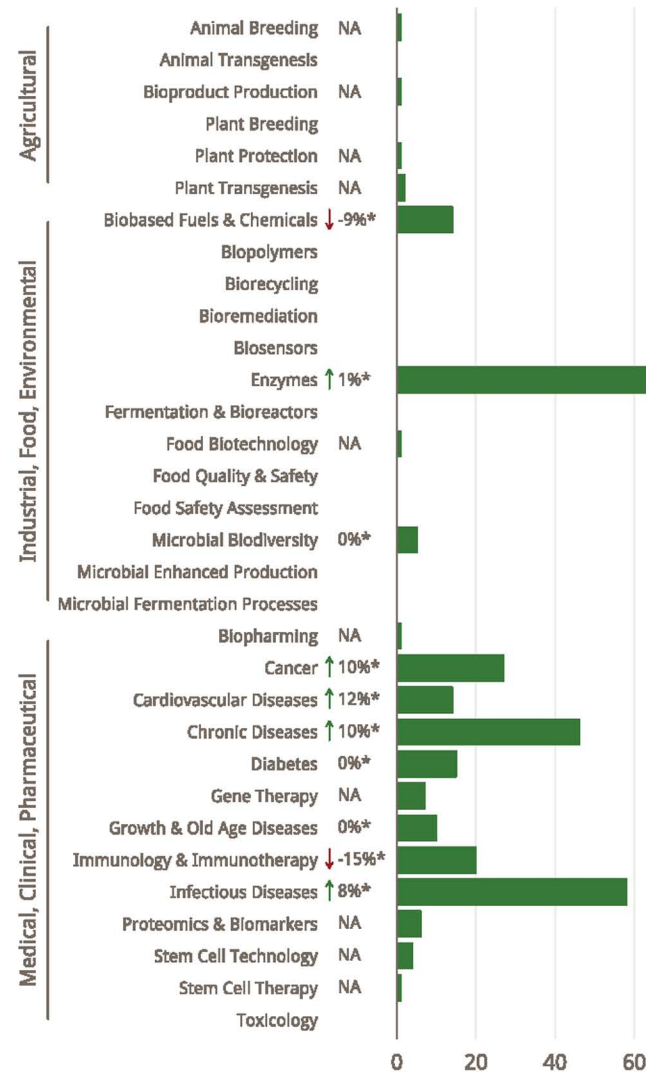
## Patent Output



CAGR: 0%\*



## Biotechnology Subcategories



## Country/Territory

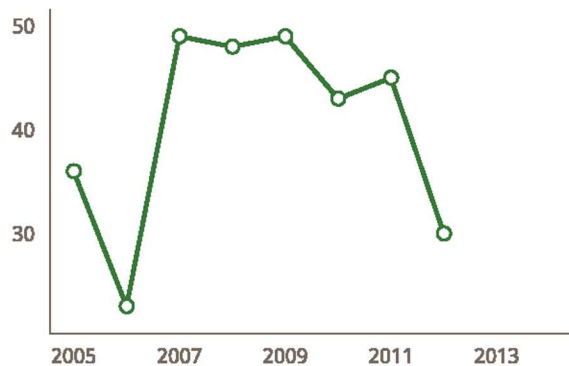


Dominican Republic

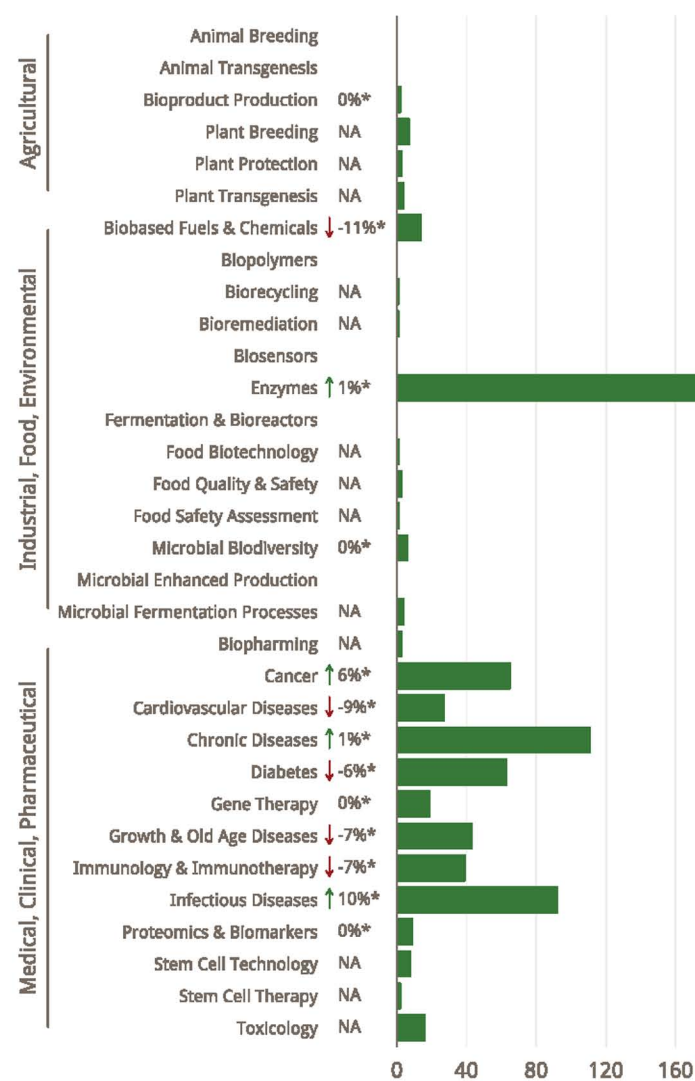
## Patent Output



CAGR: -2%\*



## Biotechnology Subcategories



## Country/Territory

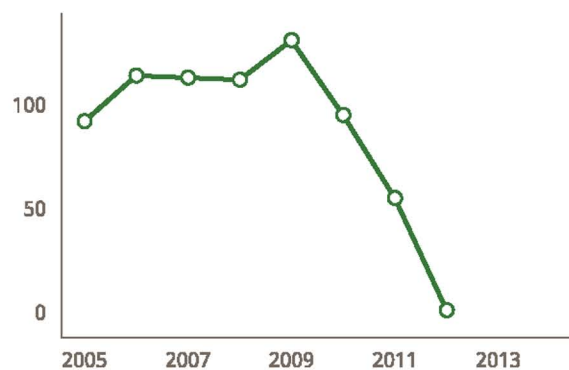


# Ecuador

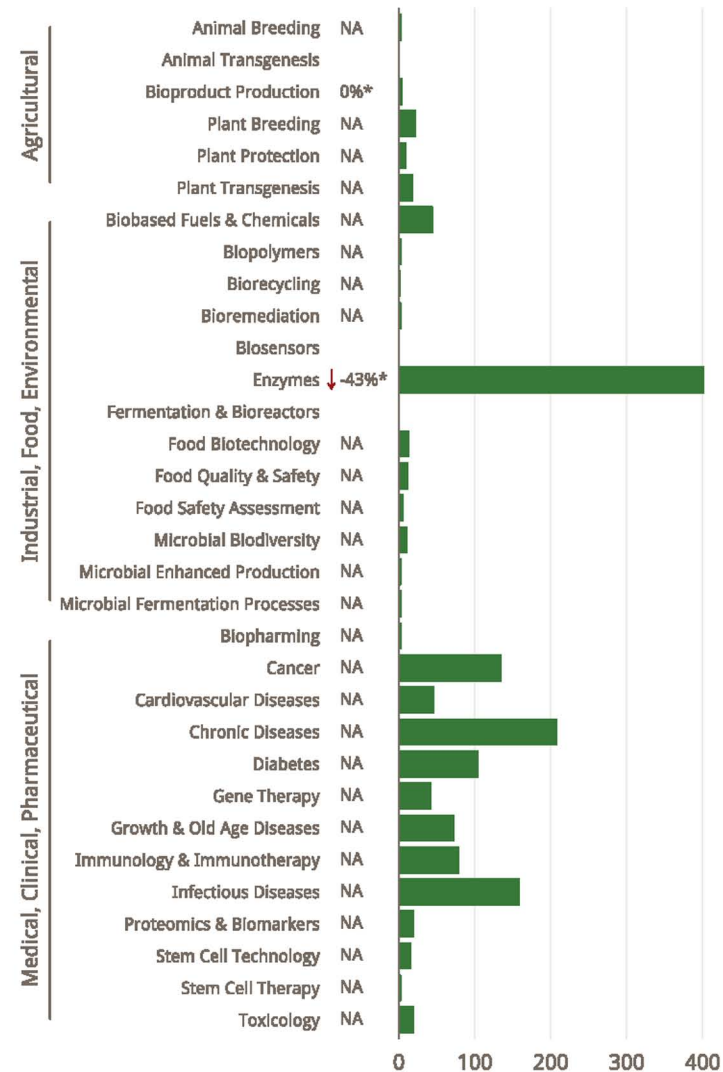
## Patent Output



CAGR: -43%\*



## Biotechnology Subcategories





## Country/Territory

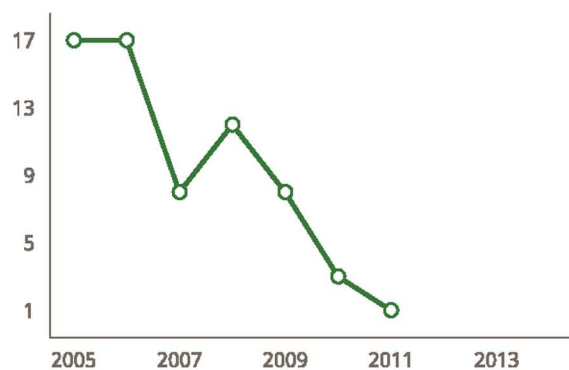


# Egypt

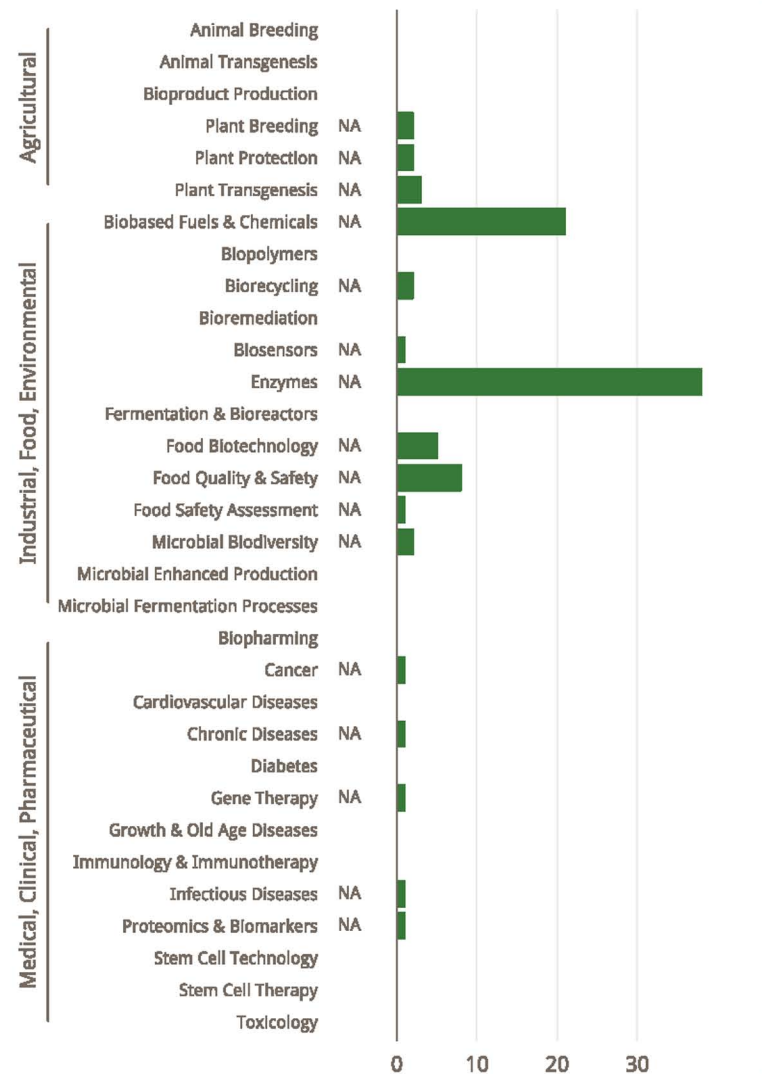
## Patent Output



CAGR: -33%\*



## Biotechnology Subcategories



## Country/Territory

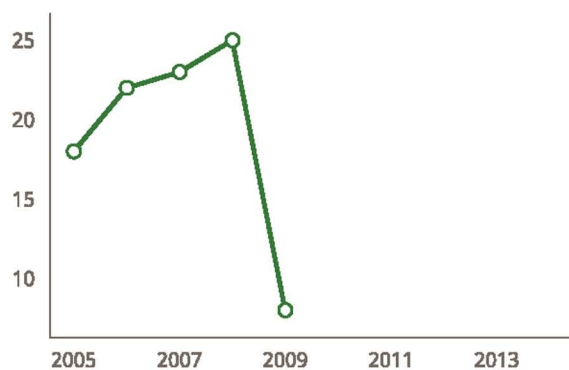


El Salvador

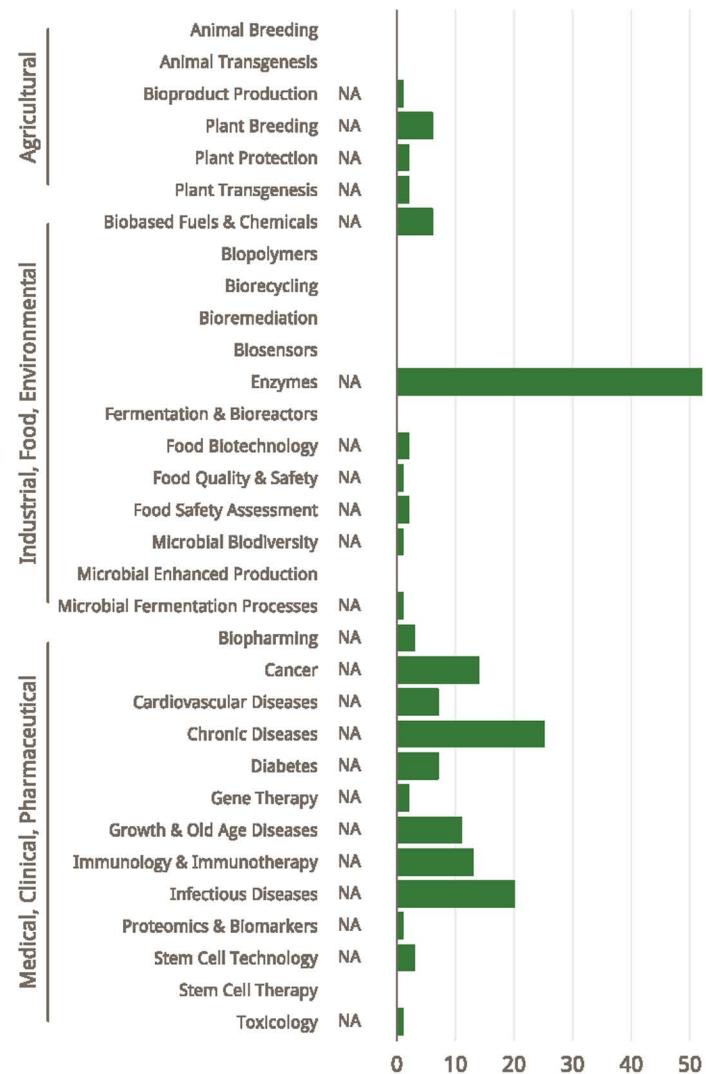
## Patent Output

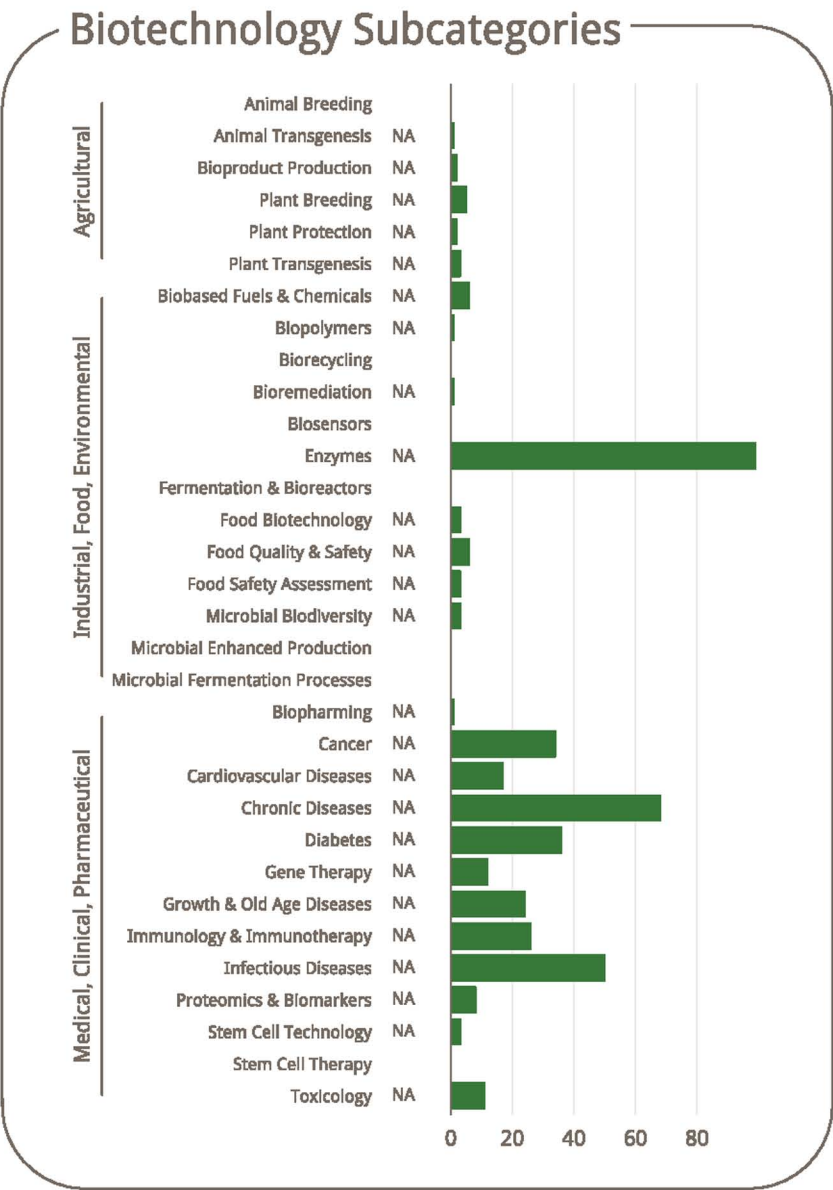
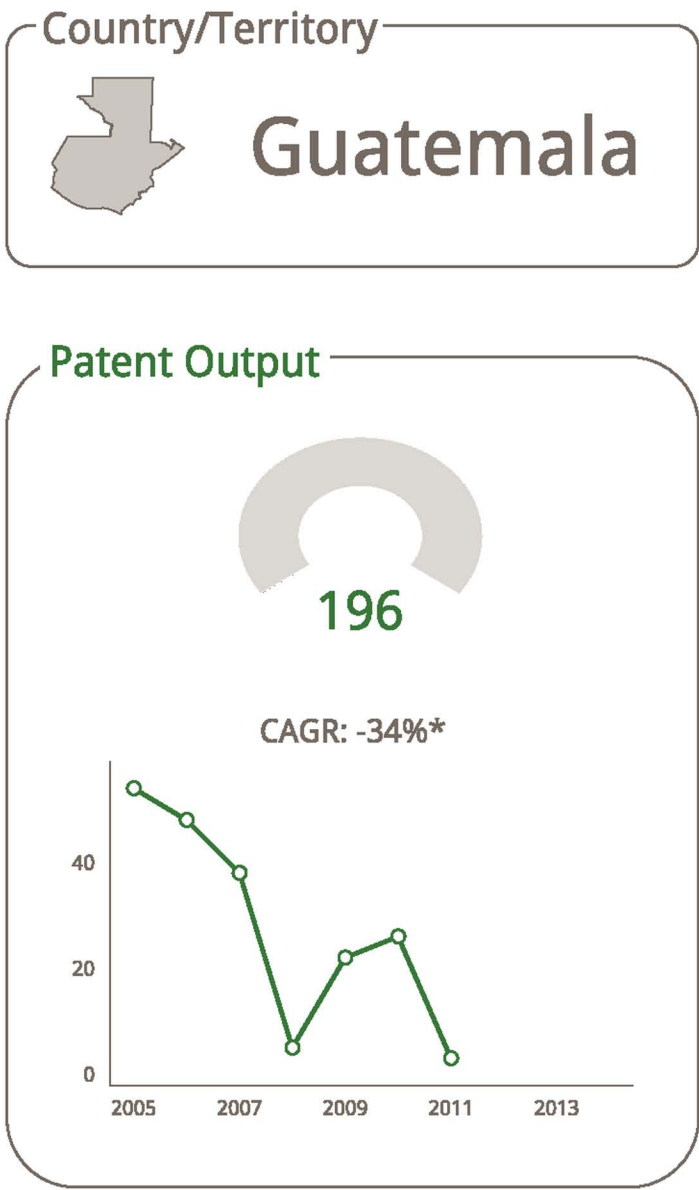


CAGR: -15%\*



## Biotechnology Subcategories





## Country/Territory

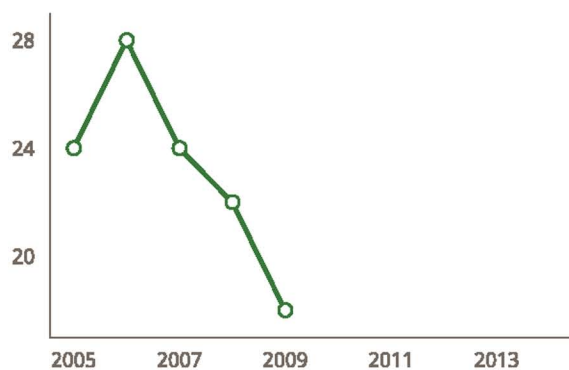


# Honduras

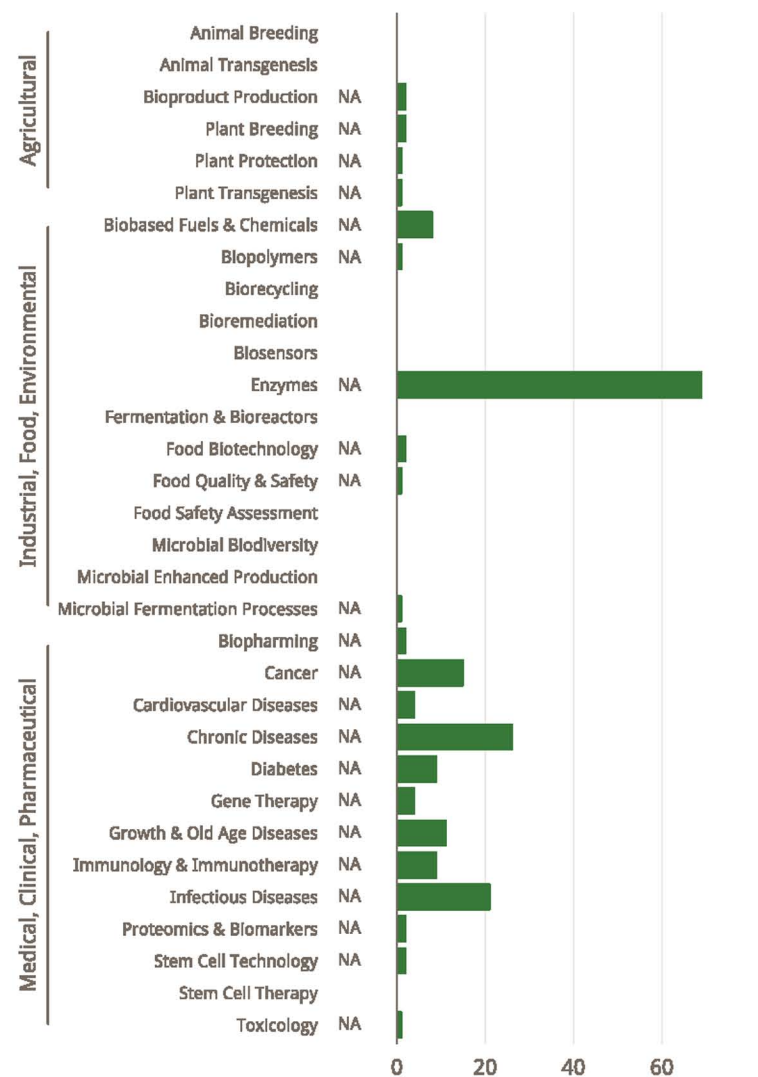
## Patent Output



CAGR: -6%\*



## Biotechnology Subcategories



## Country/Territory

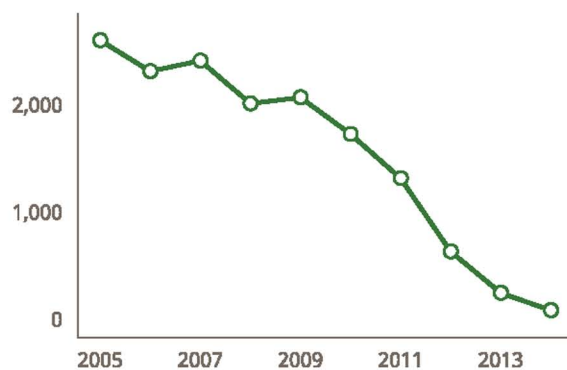


# India

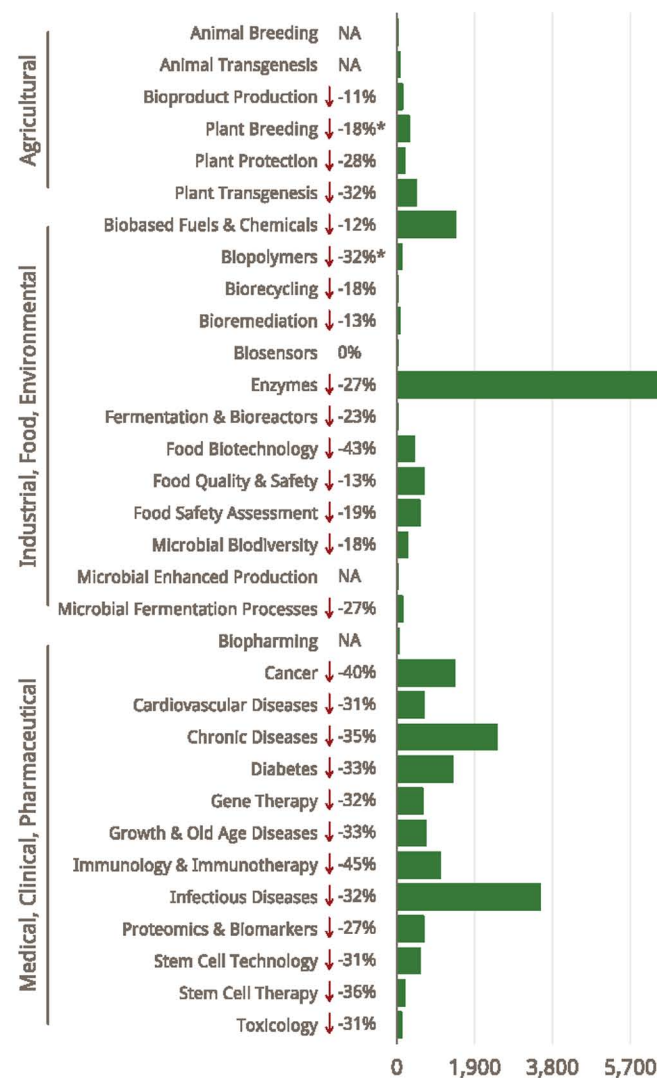
## Patent Output



CAGR: -25%



## Biotechnology Subcategories



## Country/Territory



# Indonesia

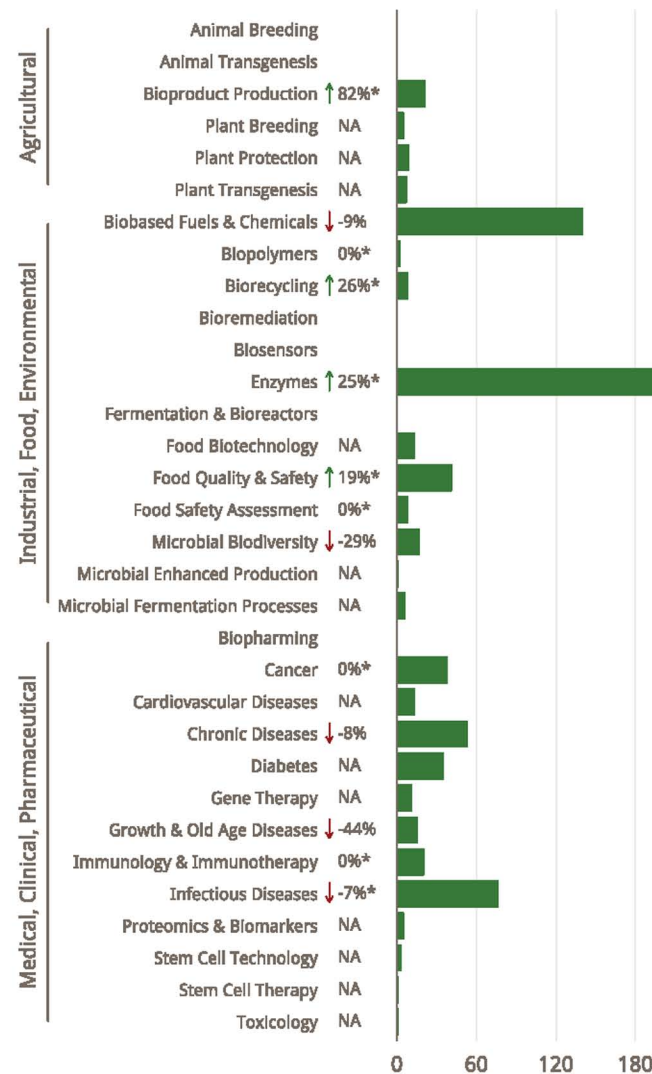
## Patent Output



CAGR: -5%



## Biotechnology Subcategories



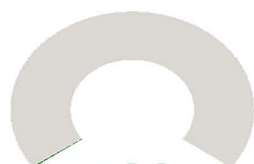


## Country/Territory

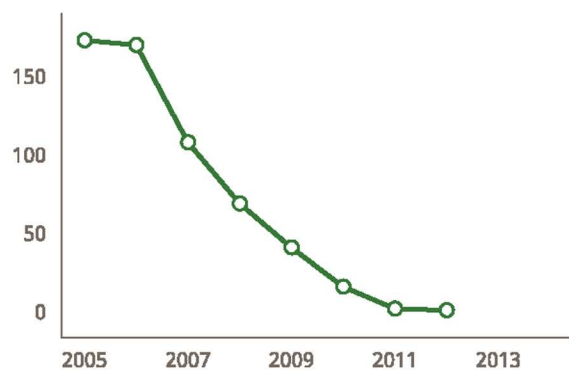


# Malaysia

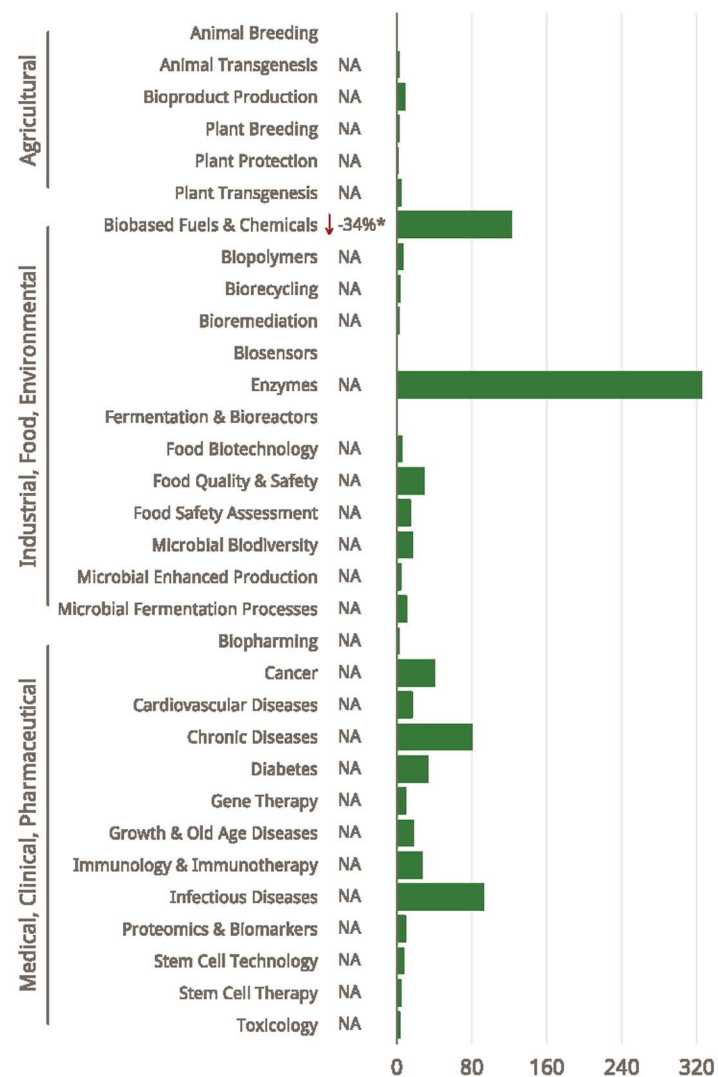
## Patent Output



CAGR: -47%\*



## Biotechnology Subcategories



## Country/Territory



# Mexico

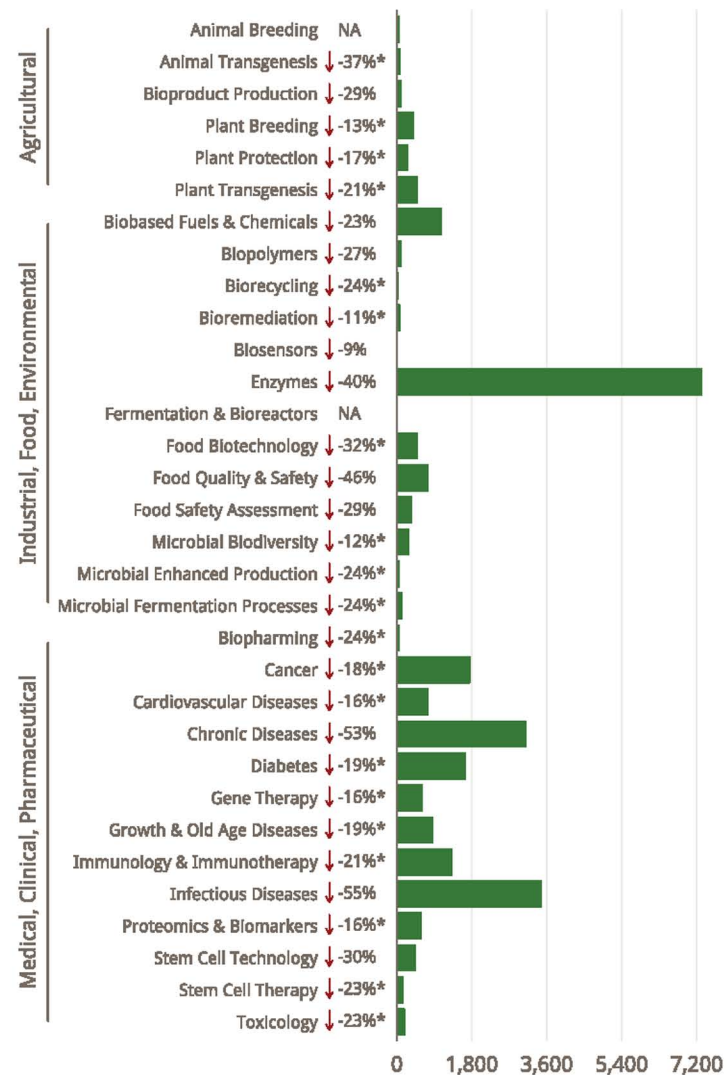
## Patent Output

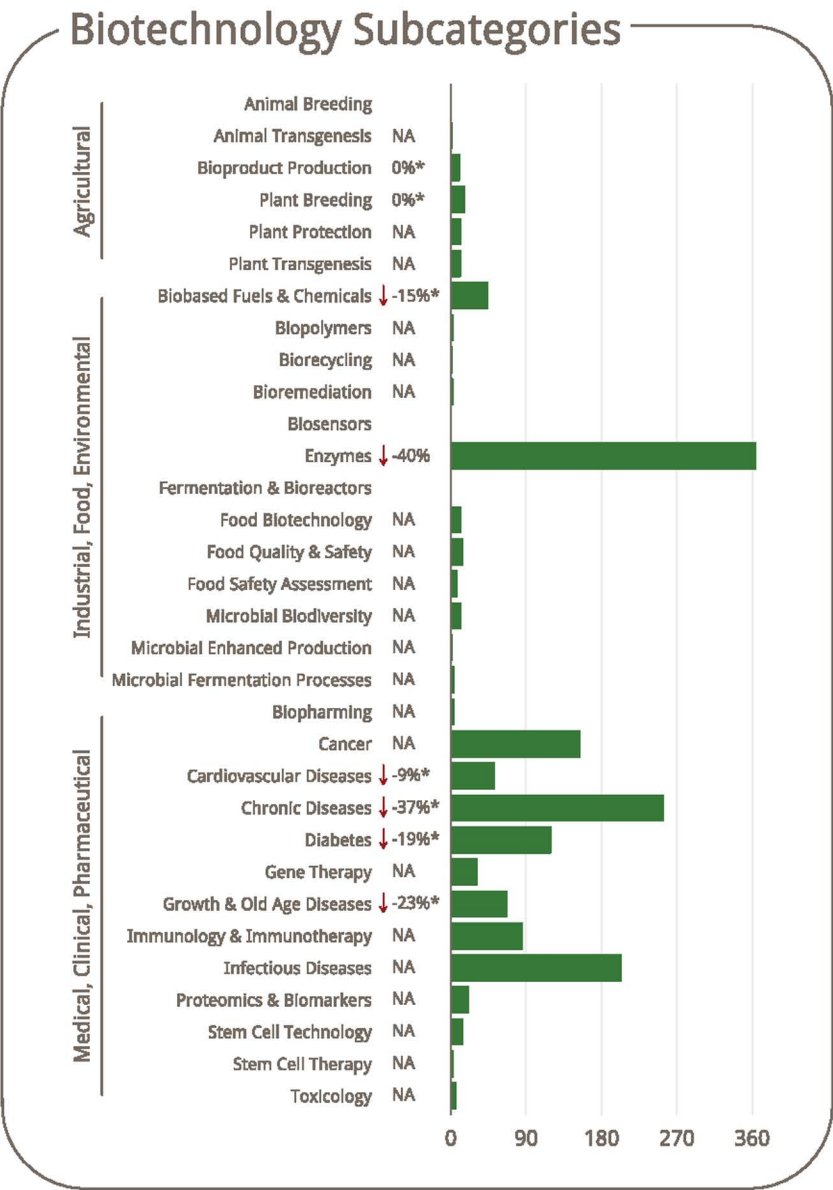
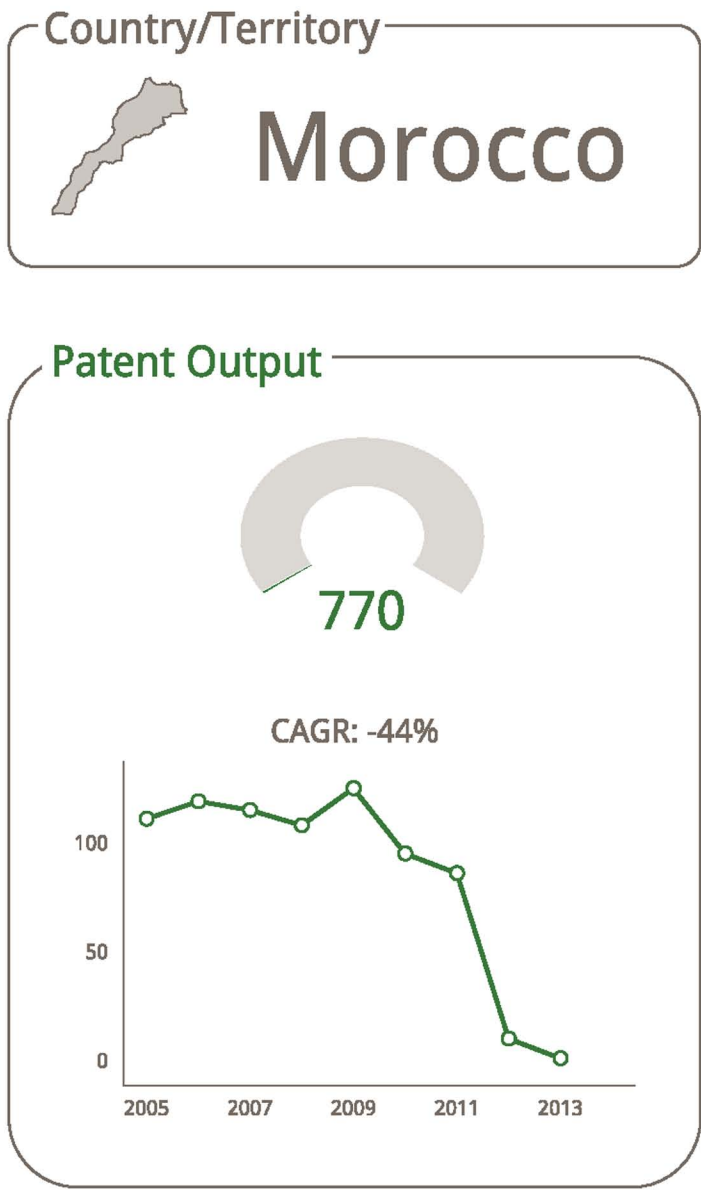


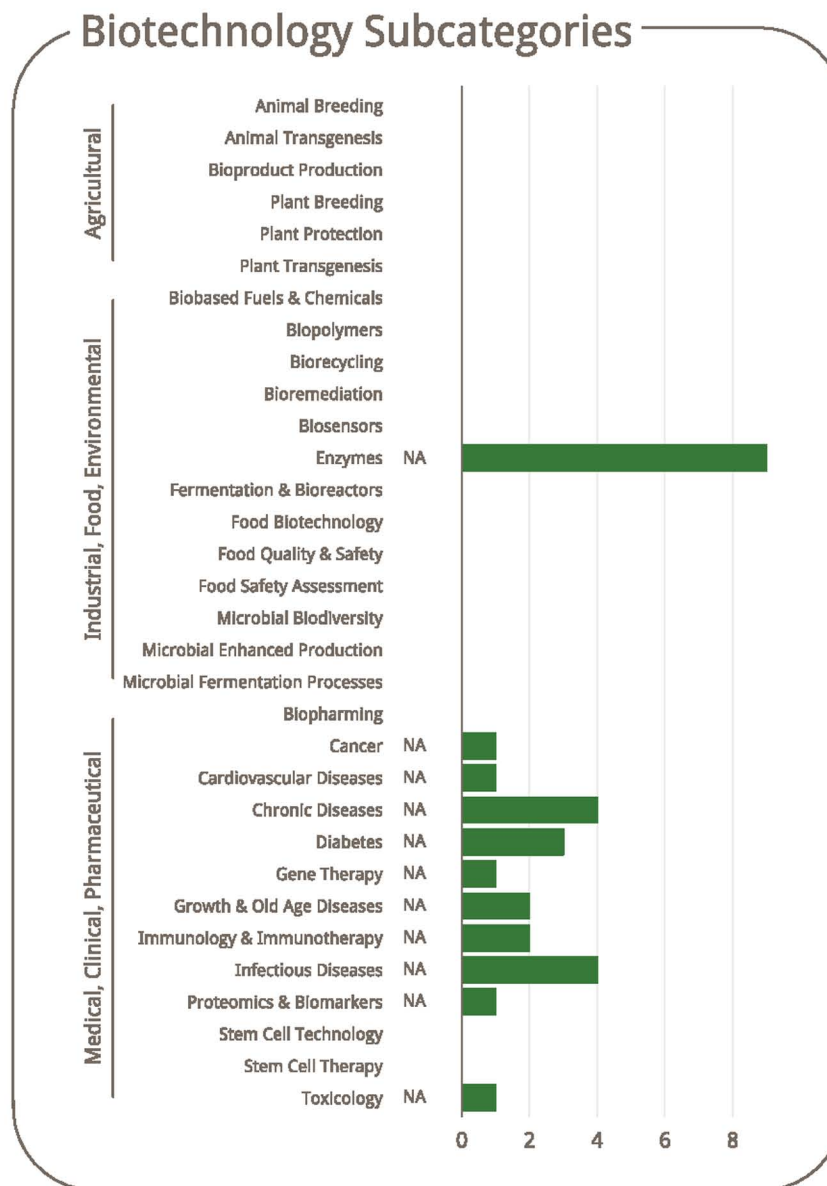
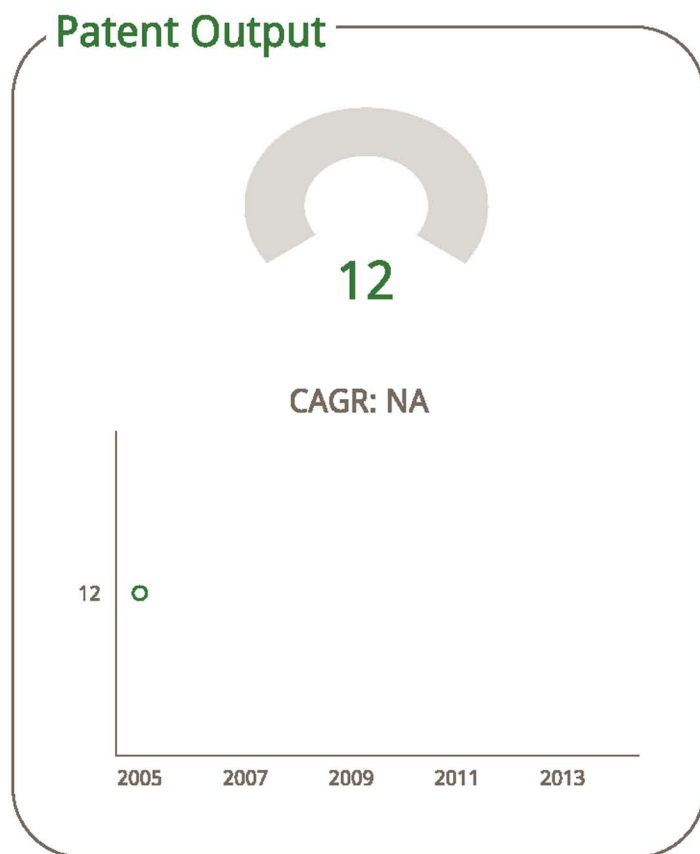
CAGR: -40%



## Biotechnology Subcategories







## Country/Territory

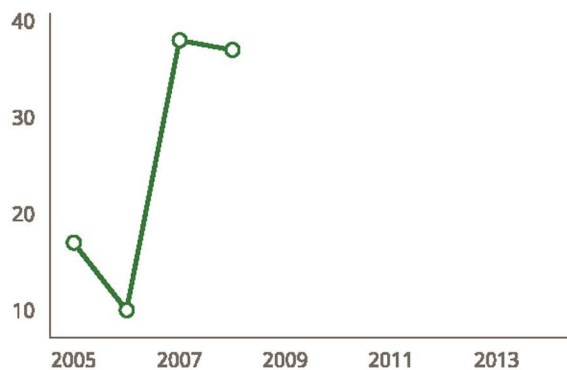


# Panama

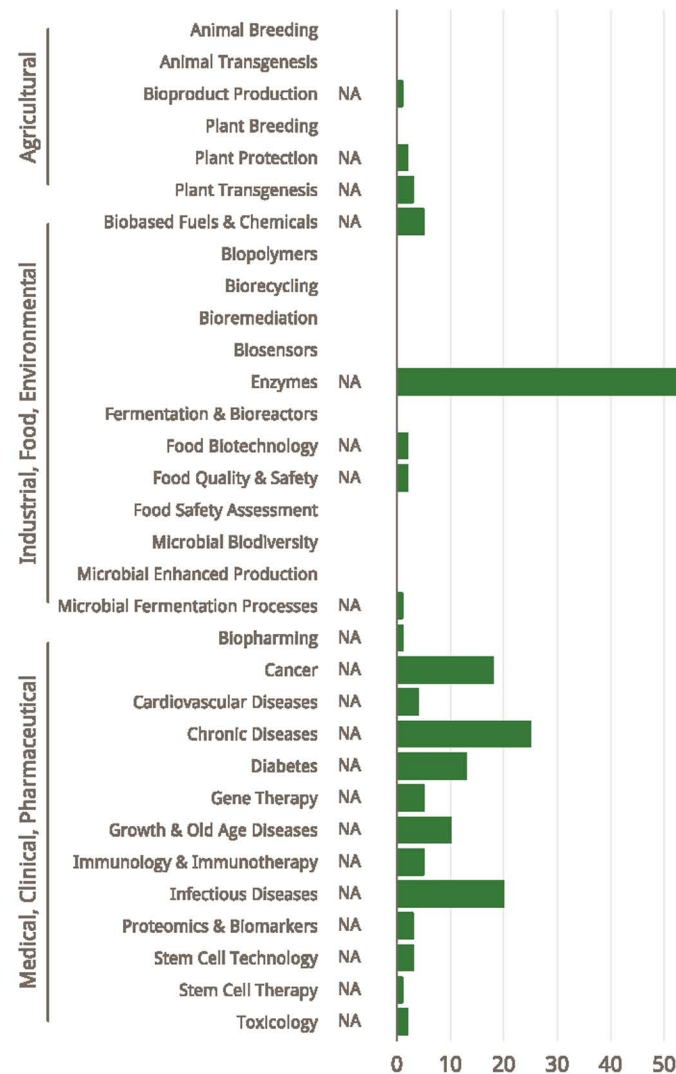
## Patent Output



CAGR: 21%\*



## Biotechnology Subcategories



## Country/Territory



Peru

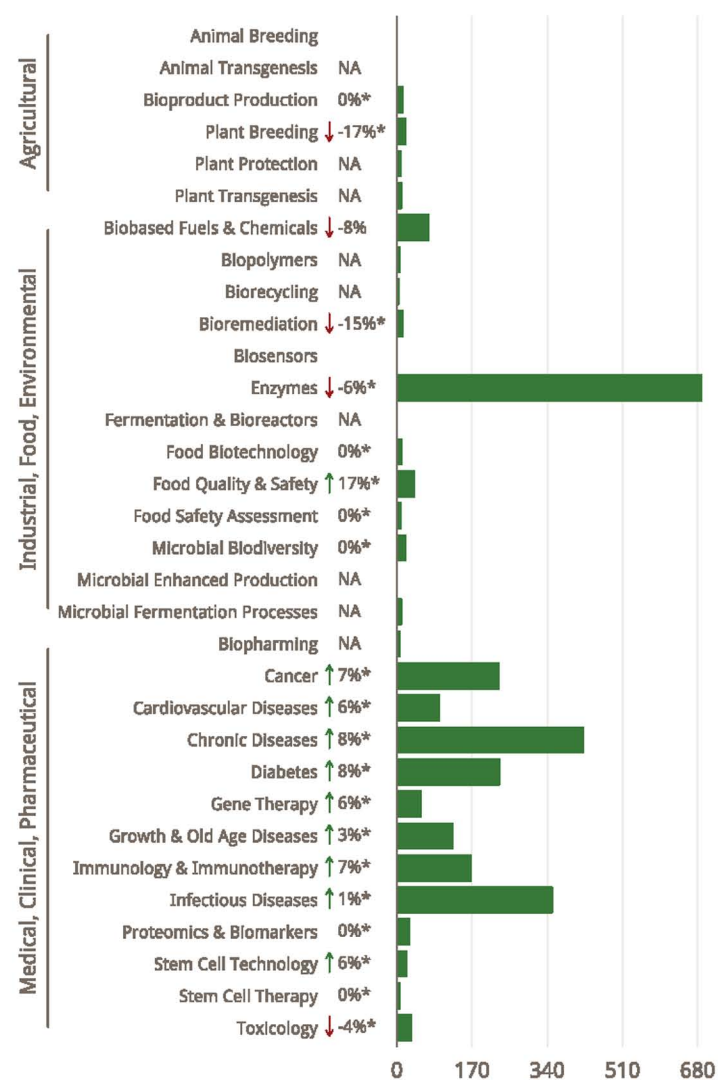
## Patent Output



CAGR: -45%



## Biotechnology Subcategories





## Country/Territory

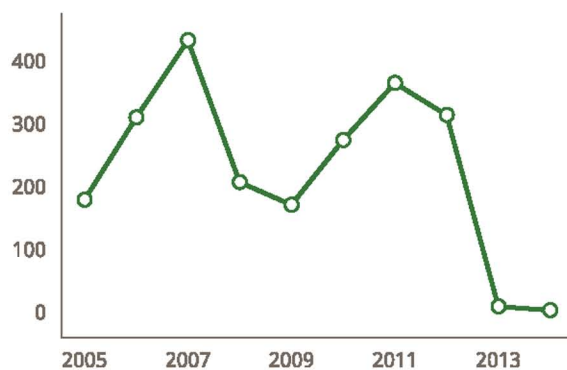


# Philippines

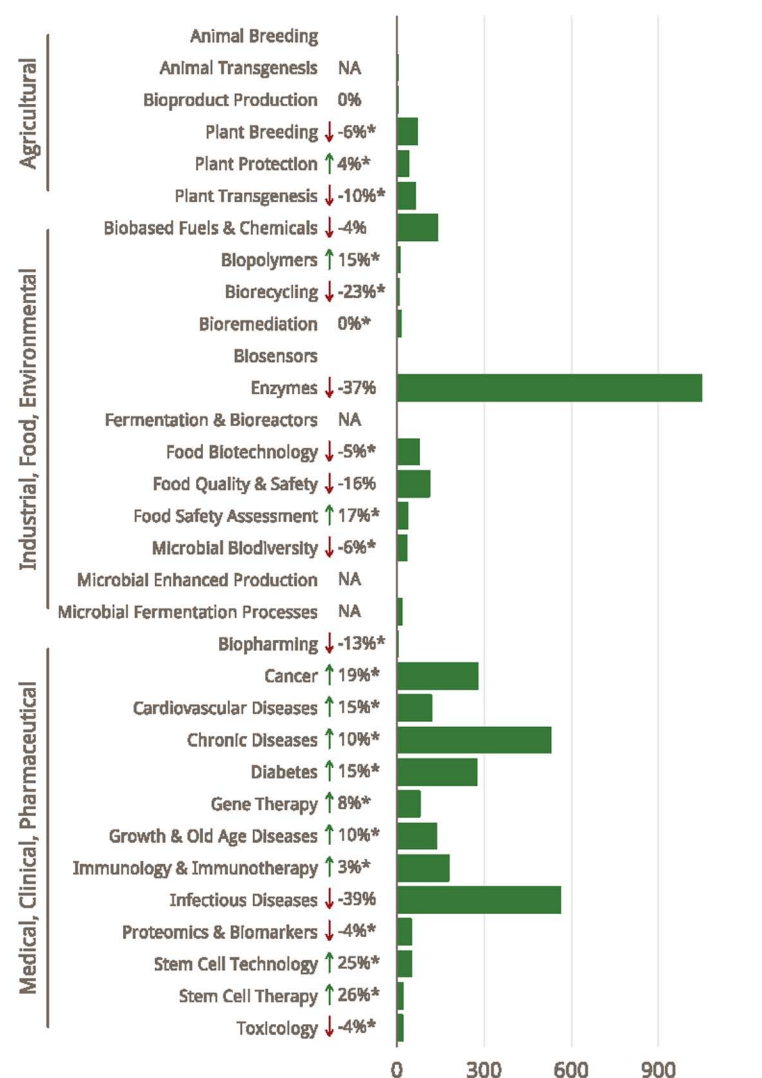
## Patent Output



CAGR: -31%



## Biotechnology Subcategories



## Country/Territory

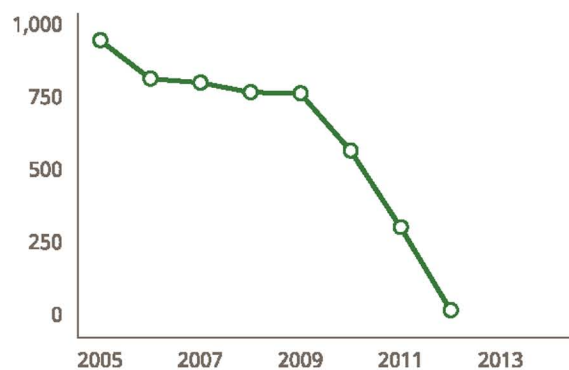


South Africa

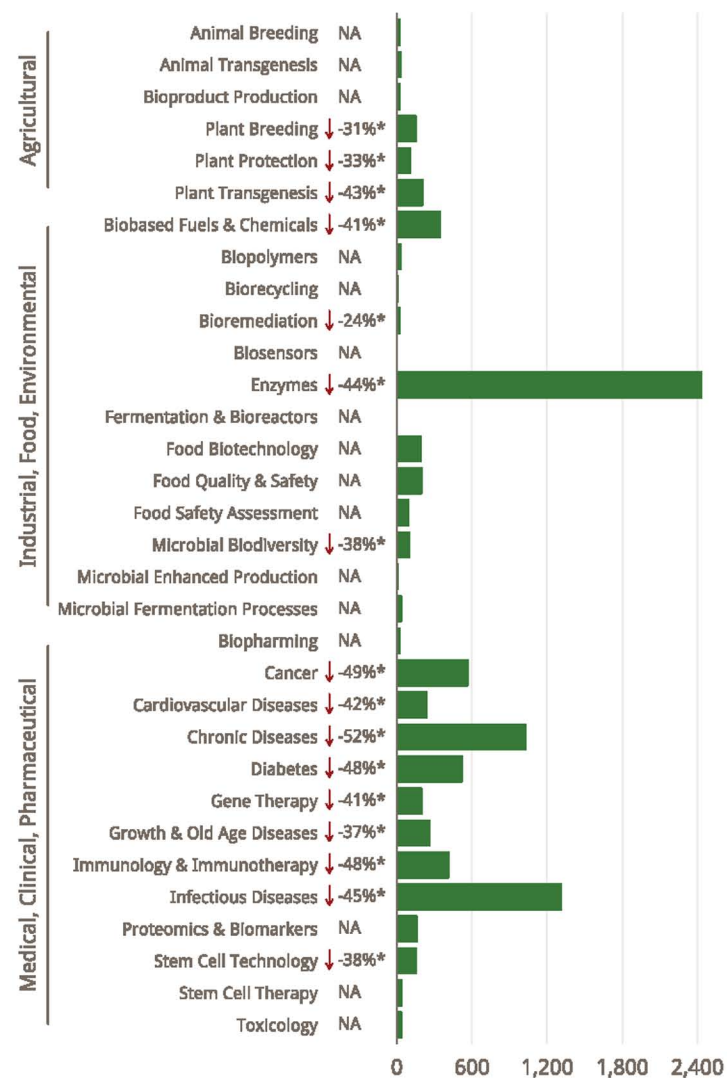
## Patent Output



CAGR: -41%\*



## Biotechnology Subcategories



## Country/Territory



# Taiwan

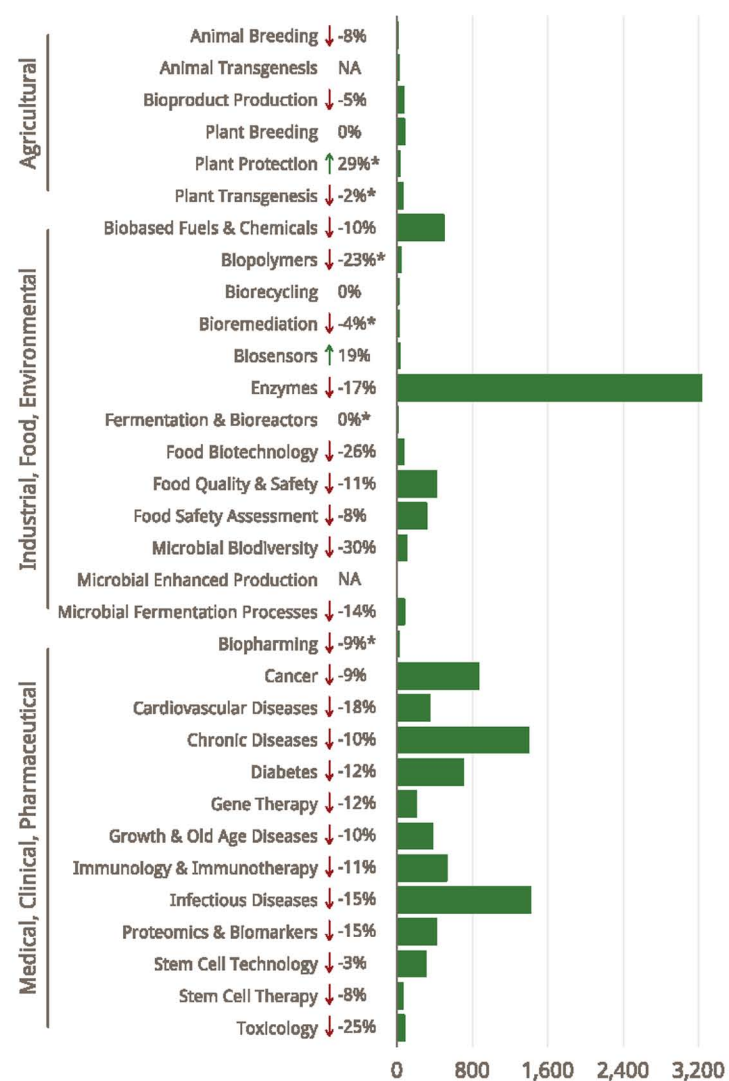
## Patent Output

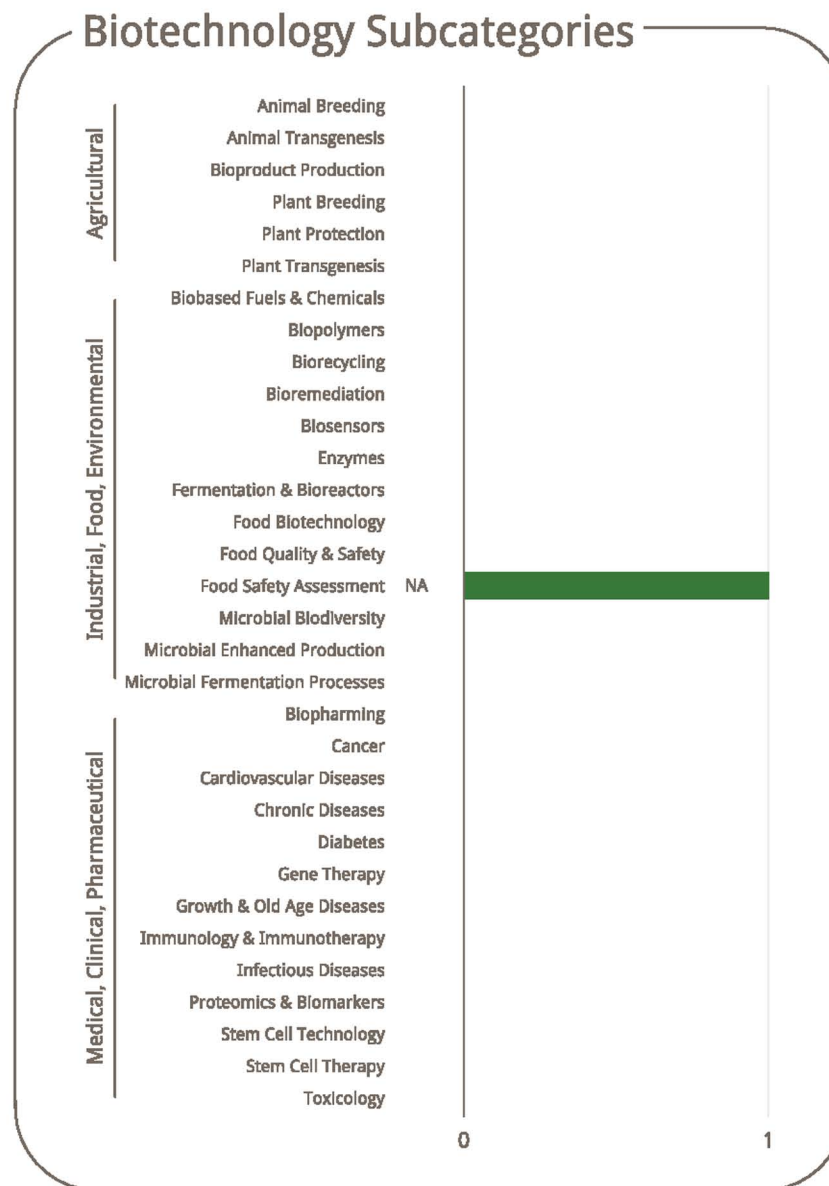
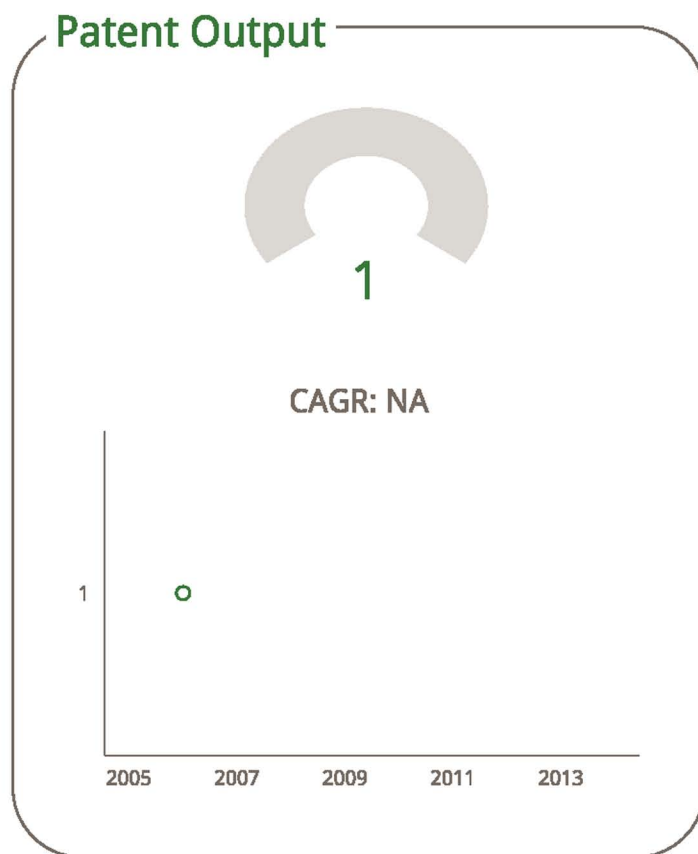


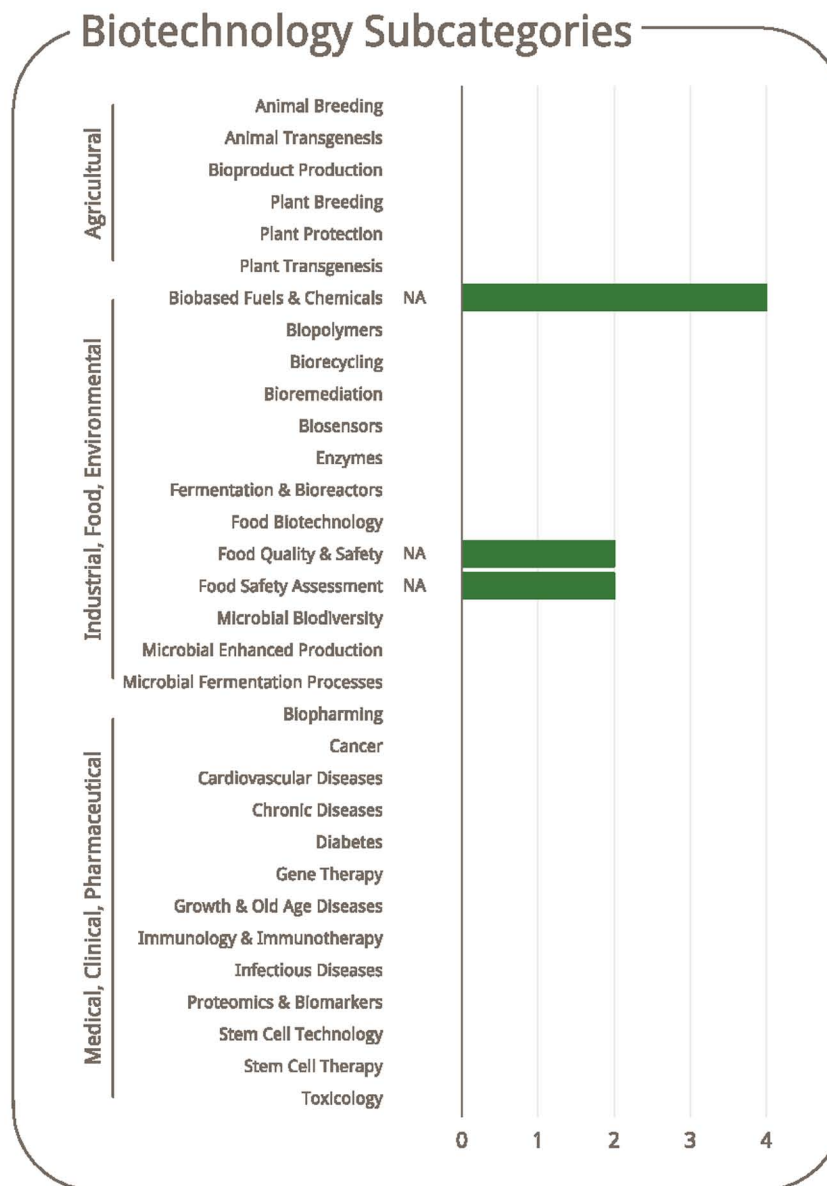
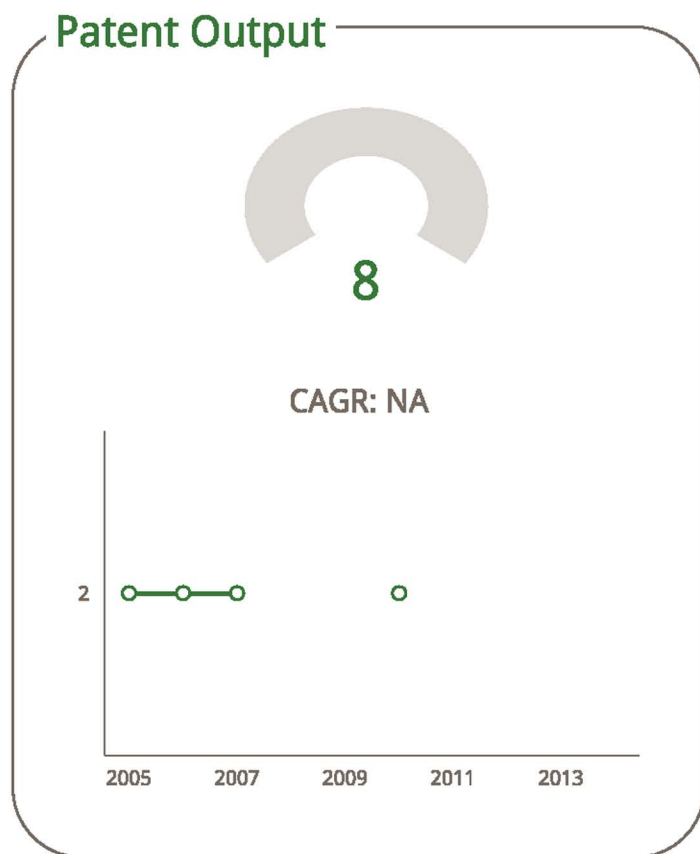
CAGR: -13%



## Biotechnology Subcategories







## Country/Territory

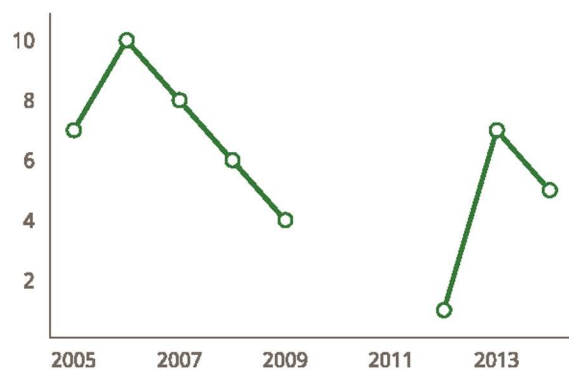


# Turkey

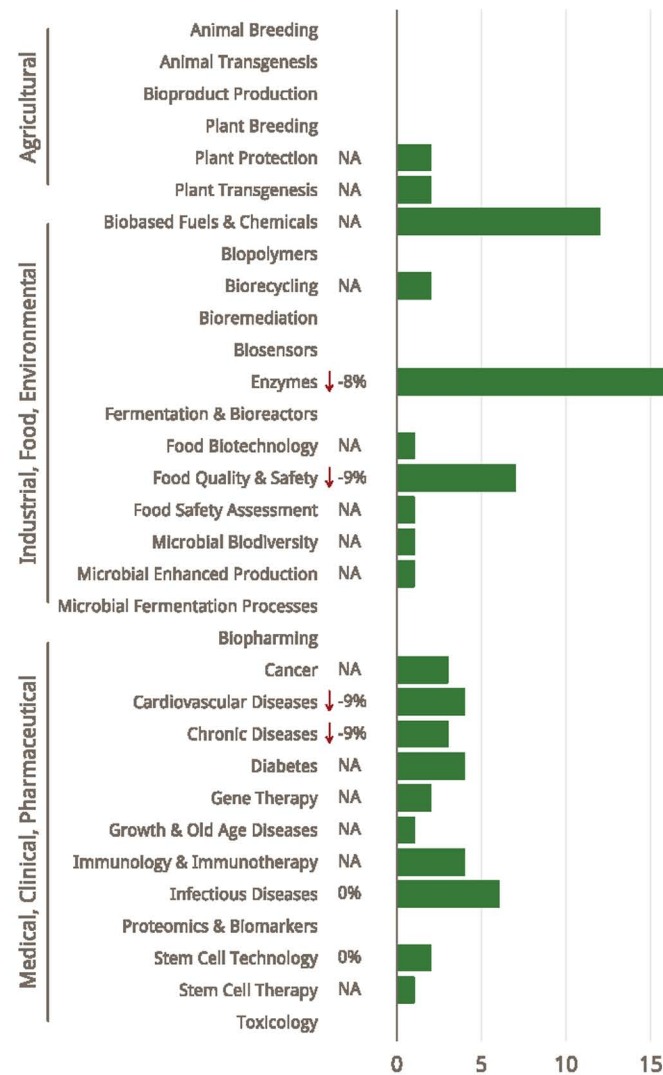
## Patent Output



CAGR: NA



## Biotechnology Subcategories





## Country/Territory



# Uruguay

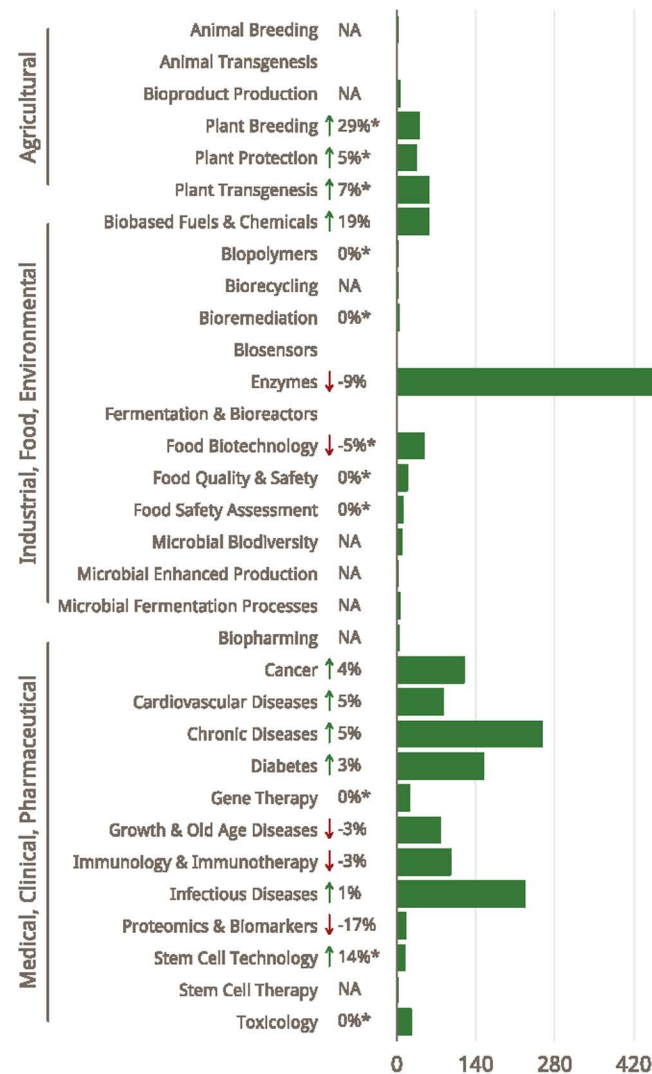
## Patent Output

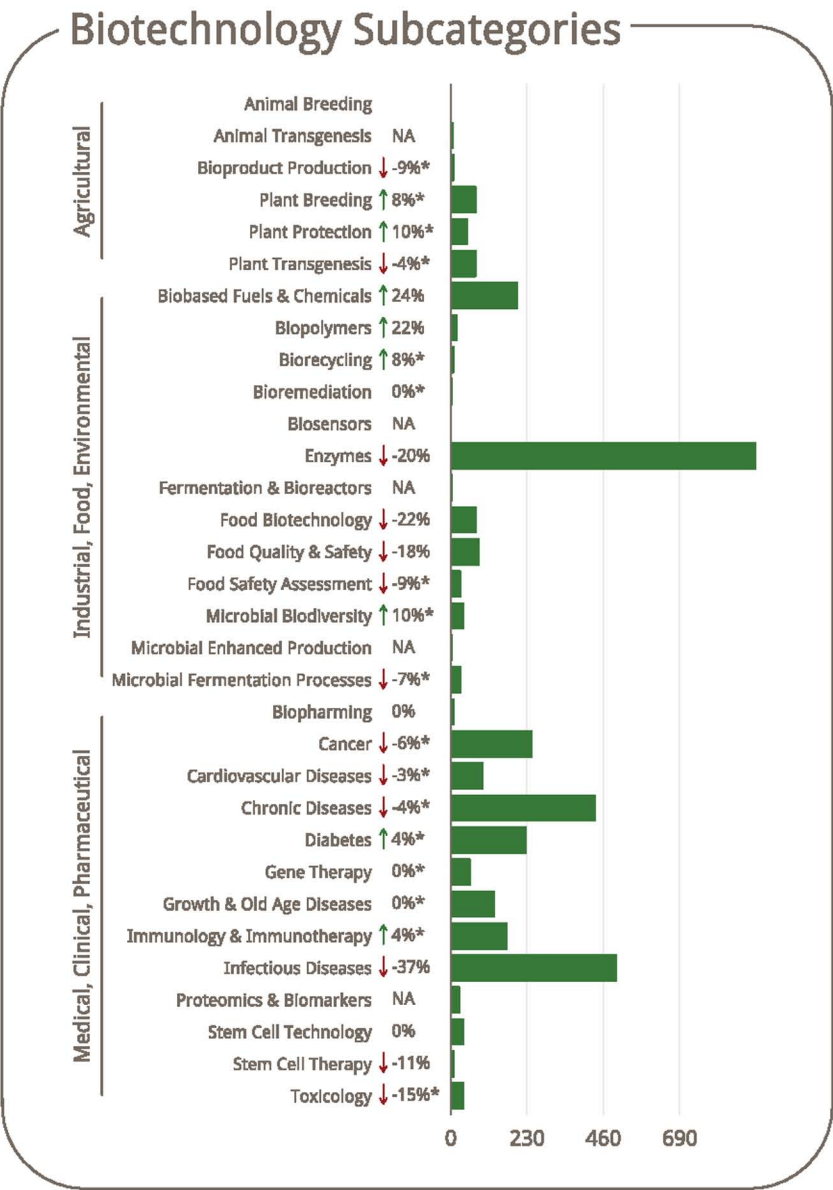
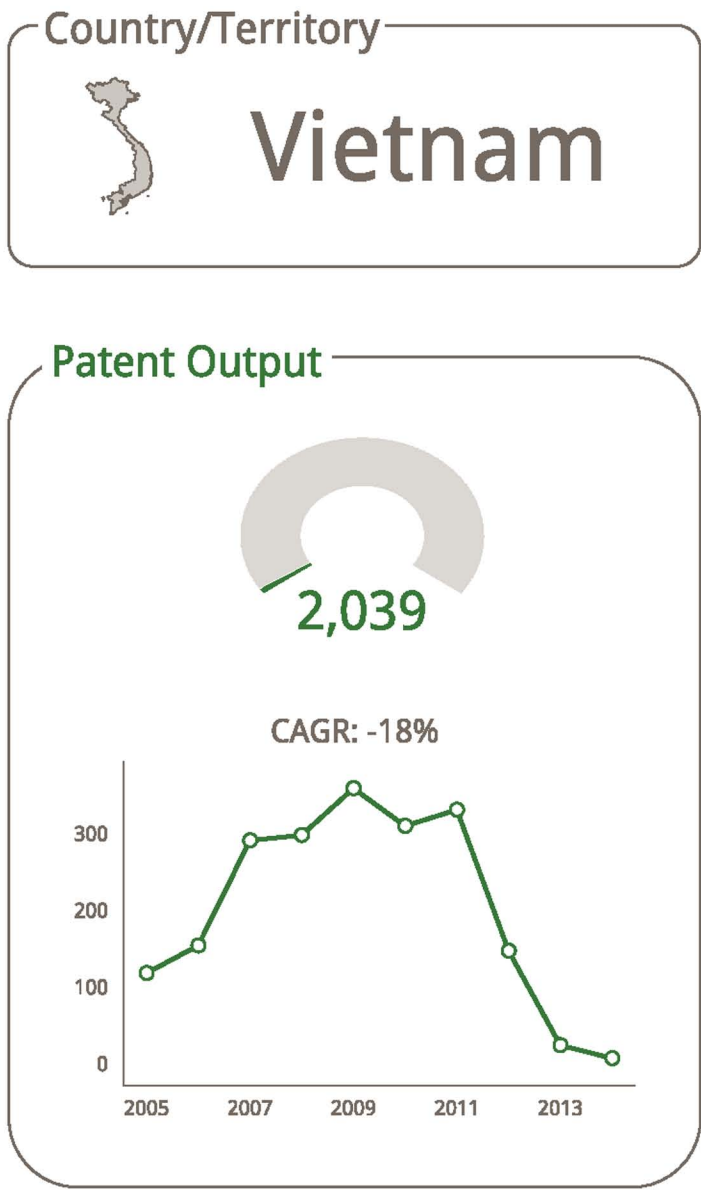


CAGR: -5%



## Biotechnology Subcategories





## Annex 1 Countries/territories covered by TWAS regional offices

TWAS Regional Office	Country/Territory
Alexandria, Egypt TWAS Arab Regional Office	Algeria
	Bahrain
	Comoro Islands
	Djibouti
	Egypt
	Iraq
	Jordan
	Kuwait
	Lebanon
	Libya
	Mauritania
	Morocco
	Oman, Sultanate of
	Palestinian Aut. Terr.
	Qatar
	Saudi Arabia
	Somalia
	Republic of South Sudan
	Sudan
	Syria
	Tunisia
	United Arab Emirates
	Yemen
Bangalore, India TWAS Regional Office for Central and South Asia	Afghanistan
	Azerbaijan
	Bangladesh
	India
	Iran, Isl. Rep.
	Kazakhstan
	Kyrgyzstan
	Maldives
	Nepal
	Pakistan
	Sri Lanka
	Tajikistan
	Turkey
	Turkmenistan
	Uzbekistan

TWAS Regional Office	Country/Territory
Nairobi, Kenya	Angola
	Benin
TWAS Regional Office for Sub-Saharan Africa	Botswana
	Burkina Faso
	Burundi
	Cameroon
	Cape Verde Islands
	Central African Rep.
	Chad
	Congo, Dem. Rep. (formerly Zaire)
	Congo, Rep.
	Cote d'Ivoire
	Eritrea
	Ethiopia
	Gabon
	Gambia
	Ghana
	Guinea
	Guinea Equatorial/Equatorial Guinea
	Guinea-Bissau
	Kenya
	Lesotho
	Liberia
	Madagascar
	Malawi
	Mali
	Mauritius
	Mozambique
	Namibia
	Niger
	Nigeria
	Rwanda
	Sao Tome and Principe
	Senegal
	Seychelles
	Sierra Leone
	South Africa
	Swaziland
	Tanzania
	Togo
	Uganda
	Zambia
	Zimbabwe

TWAS Regional Office	Country/Territory
Rio de Janeiro, Brazil	Antigua & Barbuda
	Argentina
TWAS Regional Office for Latin America and the Caribbean	Aruba
	Bahamas
	Barbados, W.I.
	Belize
	Bolivia
	Brazil
	Chile
	Colombia
	Costa Rica
	Cuba
	Dominican Rep.
	Ecuador
	El Salvador
	Falkland Islands
	Grenada
	Guatemala
	Guyana
	Haiti
	Honduras
	Jamaica, W.I.
	Mexico
	Nicaragua
	Panama
	Paraguay
	Peru
	Saint Vincent and the Grenadines
	Suriname
	Trinidad & Tobago
	Uruguay
	Venezuela

TWAS Regional Office	Country/Territory
Beijing, China	Bhutan
	Brunei
TWAS Regional Office for East and South-East Asia and the Pacific	Cambodia
	China
	Fiji
	Hong Kong
	Indonesia
	Kiribati
	Korea, DPR
	Lao PDR
	Macau
	Malaysia
	Micronesia, Federal State of
	Mongolia
	Myanmar
	Palau
	Papua New Guinea
	Philippines
	Samoa
	Solomon Islands
	Taiwan, China
	Thailand
	Timor-Leste
	Tonga
	Tuvalu
	Vanuatu
	Vietnam
	Western Samoa



## Annex 2 Scientific Lagging Countries

Afghanistan	Gambia	Niger
Angola	Ghana	Nigeria
Azerbaijan	Guatemala	Papua New Guinea
Bangladesh	Guinea	Paraguay
Belize	Guinea-Bissau	Peru
Benin	Guyana	Rwanda
Bhutan	Haiti	Samoa
Botswana	Honduras	São Tome and Principe
Burkina Faso	Indonesia	Senegal
Burundi	Jamaica	Sierra Leone
Cambodia	Kenya	Solomon Islands
Cameroon	Kiribati	Somalia
Cape Verde	Korea DPR	South Sudan
Central African Republic	Kyrgyzstan	Sudan
Chad	Lao PDR	Suriname
Comoros Islands	Lesotho	Tajikistan
Congo, Dem Rep	Liberia	Tanzania
Congo, Rep	Madagascar	Timor-Leste
Côte d'Ivoire	Malawi	Togo
Djibouti	Maldives	Tuvalu
Dominican Republic	Mali	Uganda
Ecuador	Mauritania	Uzbekistan
El Salvador	Mongolia	Vanuatu
Equatorial Guinea	Mozambique	Vietnam
Eritrea	Myanmar	Yemen
Ethiopia	Nepal	Zambia
Gabon	Nicaragua	Zimbabwe

## Annex 3 Keywords used in this project

Technology field	Subcategories	Search Concept
Medical, clinical, pharmaceutical Biotechnology	Stem Cell Therapy	(stem / pluripotent / multipotent) cell (stem / cell) therapy
	Stem Cell Technology	(stem / pluripotent / multipotent) cell (stem / cell) therapy
	Proteomics and Biomarkers	[Proteomic OR proteome (study / analysis / screen / high throughput screen / large scale screen)] [protein OR peptide] in proximity with [study OR analysis OR screen OR high throughput screen OR large scale screen] [protein OR peptide] in proximity with [structure OR posttranslational modification OR phosphorylation OR glycosylation OR acetylation OR methylation OR nitrosylation] [protein OR peptide] in proximity with [interaction OR sequencing OR mass fingerprint OR array OR molecular display OR affinity chromatography OR mass spectrometry OR NMR OR NMR spectroscopy OR nuclear magnetic resonance OR de novo sequencing OR X-ray diffraction] [protein-protein interaction OR immune precipitation OR immune precipitation OR pull down assay OR yeast two hybrid OR alanine scan OR bimolecular fluorescence complementation OR enzyme kinetic OR biochip OR MALDI-MS OR ESI OR biomarker OR biological marker]
	Immunology and Immunotherapy	[Genetic immunization OR genetic immunotherapy OR immunotherapy OR immunostimulatory] [(induce / enhance / suppress) immune response] in proximity with [immunomodulator OR interleukin OR cytokine OR interferon OR chemokine OR glucan OR glucans OR granulocyte colony-stimulating factor OR G-CSF OR Tcell OR T cell OR immune recovery OR gene transfer OR antisense OR oligodeoxynucleotide] [immunologic OR immune OR autoimmune] in proximity with [disease OR disorder OR deficiency]
	Gene Therapy	Gene therapy
	Toxicology	[disease OR poisoning OR infection OR shock OR syndrome] in proximity with [toxic OR toxin OR food OR food-borne] [toxic] in proximity with [source OR reason OR agent OR substance OR chemical OR drug OR environment]

<b>Biopharming</b>	<p>[produce OR generate OR prepare OR synthesis OR coding] in proximity with [pharma OR drug OR therapeutic OR monoclonal antibody OR vaccines OR actives]</p> <p>[transgenic OR trans-genic OR genetic (modify / alter) OR [bio-pharming OR biopharming OR molecular (pharming / farming)]</p> <p>[transgenic OR trans-genic OR genetic (modify OR alter) OR recombinant] in proximity with [mice OR mouse OR rat OR rats]</p>
<b>Chronic Diseases</b>	<p>[chronic (respiratory / pulmonary / lung / pneumonia) (disease / illness / disorder) OR asthma OR bronchiectasis OR chronic obstructive pulmonary OR bronchitis OR emphysema OR chronic rhino sinusitis OR hypersensitive pneumonitis OR lung cancer OR neoplasm respiratory organ OR lung fibrosis OR chronic pleural OR pneumoconiosis OR pulmonary eosinophilia OR pulmonary heart disease OR pulmonary hypertension OR rhinitis OR sarcoidosis OR sleep apnea OR chronic kidney disease OR CKD OR renal (disease / disorder) OR ESRD OR chronic kidney failure OR CKF OR chronic renal failure OR CRF OR obesity OR obese OR body fat OR (chronic [(Gastrointestinal disorder) OR (GI disorder) OR (digest disorder)]) OR [chronic (Gastrointestinal disease / GI disease / digest disease)] OR pancreatitis OR gastro esophageal reflux OR GERD OR chronic acid reflux OR Barrett esophagus OR Barrett OR dysphagia OR gastro esophageal mobility OR Crohn OR ulcerative colitis OR gastritis OR hepatitis C OR chronic liver disease OR anemia OR anaemia OR anemic OR antianemic OR [decrease (red blood cell)] OR hemorrhage OR hemolysis OR haemolysis OR [ineffective (hematopoiesis / hematopoiesis)]</p>
<b>Growth and Old Age Diseases</b>	<p>[Parkinson OR [(idiopathic OR primary) parkinsonism] OR hypokinetic rigid syndrome OR paralysis agitans OR Alzheimer OR paralysis OR paraplegia OR quadriplegia OR muscle function loss OR muscle nerve damage OR muscle neural dysfunction]</p> <p>[hypercortisolism OR hyperpituitarism OR hypopituitarism OR hyperthyroidism OR hypothyroidism OR [hormone (disorder / deficiency / disease) OR genetic disorder OR genetic disease OR genetic abnormality OR genetic alteration OR genetic defect OR gene alteration OR gene defect OR mutation OR point mutation]</p> <p>[acromegaly OR acromegalia OR gigantism OR giantism OR dwarfism OR dwarfness OR microsomia OR overgrowth OR stunted growth OR growth arrest OR rapid growth OR [growth (deficiency / abnormality / disorder / disease)] OR achandroplasia OR pseudoachandroplasia OR laron OR laron syndrome OR spondyloepiphyseal dysplasia OR diastrophic dysplasia OR hypochondroplasia OR Noonan OR Noonan syndrome OR primordial dwarfism OR intrauterine dwarfism OR Turner OR Turner syndrome OR osteogenesis imperfecta OR down syndrome OR trisomy 21]</p>

	<b>Infectious Diseases</b>	<p>[infection OR infectious disease OR transmissible disease OR communicable disease OR HIV OR hepatitis OR influenza OR tuberculosis OR malaria OR meningitis OR diarrhea OR diarrhoea OR respirator]</p> <p>[infection OR infectious disease OR transmissible disease OR communicable disease] in proximity with [virus OR viral OR bacteria OR microbe OR microorganism OR fungi OR fungal OR protozoa OR protozoan OR parasite OR sexual OR skin OR insect]</p>
	<b>Cancer</b>	<p>[cancer OR cancerous OR malignancy OR malignant OR malignant neoplasm OR anticancer OR antitumor] in proximity with [breast OR blood OR thyroid OR prostate OR lymphatic OR lymphoma OR leukemia OR leukaemia]</p> <p>[benign tumor]</p>
	<b>Cardiovascular Diseases</b>	<p>[Heart (angina / angina pectoris) OR heart failure OR hydrothorax OR hemothorax OR haemothorax]</p> <p>[lung OR pulmonary OR chest OR heart] in proximity with [congestion OR congest OR hyperemia OR hyperaemia]</p> <p>[blood OR blood flow OR bleed] in proximity with [clott OR thrombus OR platelet plug]</p> <p>[fibrin formed clot OR platelet aggregation OR activated platelet OR fibrin plug OR thrombotic occlusion OR thrombosis OR thrombotic disorder OR thromboembolism OR thrombotic risk OR anticlotting OR clotting factor OR hypertension OR high blood pressure OR elevated blood pressure OR prehypertension OR antihypertensive]</p>
	<b>Diabetes</b>	[Diabetes OR diabetic OR high blood sugar OR insulin resistance OR gestational diabetes OR antidiabetic]
<b>Agricultural Biotechnology</b>	<b>Plant Breeding</b>	<p>[Plant OR palm OR wheat OR vegetable OR crop OR tomato]</p> <p>[salt OR saline OR salinity OR drought OR heat OR cold OR stress] in proximity with [resistance OR resistant OR tolerance OR tolerant]</p>
	<b>Animal Breeding</b>	<p>[animal OR non-human animal OR veterinary]</p> <p>[gene diversity OR genetic diversity OR biodiversity OR biological diversity OR genetic variation OR adaptive variation OR gene pool OR marker gene OR polymorphic marker OR molecular marker OR polymorphism OR single nucleotide polymorphism OR SNP OR whole-genome sequence OR body fat OR lean muscle mass OR milk production]</p> <p>plants</p>

Technology field	Subcategories	Search Concept
<b>Agricultural Biotechnology (continued)</b>	<b>Bioproduct Production</b>	<p>[investigate OR identify OR analysis OR analyze OR extract OR isolate OR separate OR synthesis OR prepare OR generate OR produce OR fermentation] in proximity with [natural OR native OR biologic origin]</p> <p>[drug OR compound OR chemical composition OR vaccine OR vitamin OR antibody OR antibacterial OR antiviral OR antifungal OR antiparasitic OR anti-inflammatory OR anticancer OR antiseptic OR analgesic OR anaesthetic]</p> <p>[plant OR crop OR leaf OR flower OR bark OR seed OR wood OR bulb OR tuber OR branch OR pericarp OR root OR herb OR herbs OR rhizome OR spor OR starch OR resin OR oil OR juice OR gum OR balsam OR wax OR tar OR alge OR algae OR vegetation]</p> <p>[pharmaceutics botany OR pharmaceutical plant OR pharmacognos OR photochemistry OR ethnobotan OR ethnopharmacolog OR phytopharmaceutic OR herbal medicine OR native medicinal plant]</p> <p>[transgenic OR trans-genic OR genetic (modify / alter) OR recombinant] in proximity with [plant OR crop OR microbe OR organism OR animal OR nonhuman OR mammal OR grass OR bacteria OR yeast OR alge OR algae OR vegetation]</p>
	<b>Plant Protection</b>	<p>[Plant OR palm OR wheat OR vegetable OR crop OR tomato]</p> <p>[nematode OR insect OR virus OR viral OR fungi OR bacteria] in proximity with [resistance OR resistant OR resist OR tolerance OR tolerant OR protection OR protected]</p>
	<b>Animal Transgenesis</b>	[Transgene OR gene OR genome] [integration OR insertion OR targeted insertion OR transformation OR ES cell OR embryonic stem cell OR pronucle injection OR retroviral insertion OR knockout OR RNAi OR somatic cell mutagenes OR gene targeting OR mutagenesis OR chemical mutagenesis OR ENU OR transposon OR embryonic injection OR (shRNA / siRNA) silencing OR cell fusion OR nuclear transplant OR electroporation]
	<b>Plant Transgenesis</b>	transgene OR gene transfer OR soil bacteria OR agrobacteria OR biolistic OR gun OR BAC OR BIBAC OR bacterial artificial chromosome OR T-DNA OR transposon OR germplasm OR transfection OR transduction OR electroporation OR microinjection OR lipofection

Technology field	Subcategories	Search Concept
Industrial, Food , Environmental Biotechnology	Microbial Biodiversity	<p>[archaeobacterial OR archaea OR nanoorganism OR nanoorganism]</p> <p>[phylogenetic OR evolution biology OR evolution relationship]</p> <p>[archaeobacterial OR archaea OR nanoorganism OR nanoorganism OR extremophile OR acidophile OR alkaliphile OR anaerobe OR cryptoendolithe OR halophile OR hyperthermophile OR thermophile OR thermoacidophile OR hypolith OR lithoautotroph OR metallotollerant OR oligotrophy OR osmophile OR piezophile OR polyextremophile OR pyschrophile OR radioresistant OR radioresistance OR xerophile OR acidophilic OR alkaliphilic OR anaerobic OR extremophilic OR cryptoendolithic OR halophilic OR hyperthermophilic OR thermophilic OR thermoacidophilic OR lithoautotrophic OR metallotollerant OR oligotrophic OR osmophilic OR piezophilic OR polyextremophilic OR pyschrophilic OR radioresistant OR radioresistance OR xerophilic OR extreme environment OR hot spring OR geyser OR black smoker]</p> <p>[use OR produce OR consume OR improve OR degrade OR biodegrade OR cycle OR recycle OR decompose OR restore OR destroy OR isolate OR extract OR control OR monitor OR predict] in proximity with [industry OR food OR beverage OR metabolite OR gas OR pollutant OR enzyme OR environment OR environmental change OR nutrient OR fuel OR biopharmaceutical OR climate OR waste OR sewage OR toxic OR chemical OR ecosystem OR hydrocarbon OR carbon OR sulfur OR sulphur OR nitrogen OR peroxide OR methane OR ammonium OR (DNA / Taq of Pfu) polymerase OR anaerobic digest OR plant pest OR animal pest OR bleach OR bioleach OR pasteurization OR decontaminate OR remediation OR food packaging OR milk OR whey OR metal OR gold OR cobalt OR copper]</p> <p>[identification OR isolate OR enrich OR cultivate OR screen OR culture OR colony OR colonies OR activity OR activities OR genome sequence OR gene transfer OR gene flow OR DNA profiling OR rRNA OR ribosomal RNA OR 16S high variable region OR freeze OR dryfreez OR barcoding OR bioprospecting]</p>
	Bioremediation	<p>[Bioremediation OR phytoremediation OR bioventing OR bioleaching OR landfarming OR land farming OR bioremediation bioreactor OR composting OR bioaugmentation OR rhizofiltration OR biostimulation OR microorganism OR micro organism OR bioremediator OR plant OR genetically modified organism OR genetically engineered micro organism]</p> <p>[Remove OR degrade OR assimilate OR capture] in proximity with [pollutant OR toxic agent OR toxic substance OR toxicant OR toxin OR metal OR pesticide OR solvent OR explosive OR crude oil OR oil OR contaminant OR contaminate OR nutrient]</p>



Technology field	Subcategories	Search Concept
<b>Industrial, Food , Environmental Biotechnology (continued)</b>	<b>Microbial Enhanced Production</b>	[Microbe OR bacteria OR microorganism OR micro-organism] in proximity with [oil OR fuel OR crude oil OR (oil / fuel) (production / desulphurization / desulfurization) OR (sulphur / sulfur) impurities OR fuel-producing bacteria or Amino acid or Alcohol or Hydrogen or Antibiotics or Food processing (diary or milk or (food additive) or cheese or curd or yogurt) or (gas/biogas/methane) or surfactant or Soil (fertility/sustainability) or Crop/agricultural (Yield improvement-/Productivity) or Organic fertilizer or biocementation or organic acid]
	<b>Microbial Fermentation Processes</b>	[Microbe OR Microorganism OR bacteria OR micro-organism OR Protozoa OR Yeast OR Fungi OR Actinobacteriaceae OR mould OR GMO OR lactic acid bacteria OR LAB OR Sewage] in proximity with [ferment or Fermentation OR Fermentative OR decompose or decomposition OR Souring OR Brewing OR winemaking OR Treatment]
	<b>Food Quality and Safety</b>	[Food OR drink OR meat OR snack OR cuisine OR feed OR beverage OR dairy OR confectionery OR dessert OR milk] in proximity with [Quality OR Unadulterated OR safety OR preparation OR Storage OR Label OR labelling OR Healthy OR Hygiene OR certification OR packaging OR delivery OR handling OR cleaning OR Control OR Security OR Preservation OR Sanitation]
	<b>Food Safety Assessment</b>	[Food OR drink OR meat OR snack OR cuisine OR feed OR beverage OR dairy OR confectionery OR ingredients OR crop OR plant] in proximity with [Quality OR safety OR preparation OR storage OR label OR labelling OR hygiene OR certification OR packaging OR delivery OR handling OR cleaning OR control OR security OR preservation OR hazard OR treatment OR Adulteration] in proximity with [Control OR assessment OR analysis OR management OR supervision OR monitor OR monitoring OR detection OR concern]
	<b>Biopolymers</b>	[synthesis OR fermentation OR production OR produce OR conversion OR convert OR grow OR engineer OR genetic engineer] in proximity with [biopolymer OR natural polymer OR bio-derived polymer OR polysaccharide OR polyester OR polyamide OR bioplastic OR exopolysaccharide OR EPS OR xanthan OR alginate OR cellulose OR cyanophycin OR poly gamma-glutamic acid OR gamma PGA OR levan OR hyaluronic acid OR organic acid OR oligosaccharide OR polysaccharide OR polyhydroxyalkanoate OR polyhydroxybutyrate OR PHB OR PHBV OR Poly lactic acid OR PLA OR CPG]
	<b>Biosensors</b>	[biological (Element / molecule / material / particle) OR tissue OR organism OR organelles OR cell receptor OR enzymes OR antibody OR microbe OR microbial OR bacteria OR yeast OR virus OR algae OR immunosystems] in proximity with [sensitive OR sensor OR sensing OR transducer]

Technology field	Subcategories	Search Concept
<b>Industrial, Food , Environmental Biotechnology (continued)</b>	<b>Biosensors (continued)</b>	[detect OR measure OR report OR identify OR read OR indicate OR sense OR record] in proximity with [environ (contaminate / toxic / toxin) OR pollutant OR pollution OR impurity OR impureness OR atmosphere OR oil OR harmful OR hazard]  [biosensor OR bio-sensor OR biodetector OR bio-detector OR bioreceptor OR bio-receptor OR bioreporter OR bio-reporter] in combination with [environ (contaminate / toxic / toxin) OR pollutant OR pollution OR impurity OR impureness OR atmosphere OR oil OR harmful OR hazard]
	<b>Fermentation and Bioreactors</b>	[fermentation OR production OR produce] in proximity with [bioreactor OR biofilm OR bioreactor OR fermentation tank OR biofilm]
	<b>Biorecycling</b>	[Biorecycl OR fermentation OR biodegrade OR decomposition OR decompose OR detoxification OR detoxify] in proximity with [waste OR industrial waste OR agricultural waste OR medical waste OR medicinal waste]  [bacteria OR microbe OR microorganism OR micro-organism]
	<b>Food Biotechnology</b>	[food / nutrition / crop / animal / vet / plant / fruit / vegetable / corn / soy / cheese / milk / grain / dairy / enzyme / sweetner / amino acid / flavour / vitamins / organic acids / polyunsaturated fatty acids / carbohydrates / metabolites] in proximity with [Genetically modified / altered / engineered / manipulated or GMO or transgenic or biotechnology / breeding / Hybridisation / mutagenesis / conjugation / mutation / selection / transformation / Spheroplast fusion / Rare mating / Single-chromosome transfer]

## Annex 4 Bibliometrics and Citation analysis

Bibliometrics are about publications and their citations. The academic field emerged from 'information science' and now usually refers to the methods used to study and index texts and information.

Publications cite other publications. These citation links grow into networks, and their numbers are likely to be related to the significance or impact of the publication. The meaning of the publication is determined from keywords and content. Citation analysis and content analysis have therefore become a common part of bibliometric methodology. Historically, bibliometric methods were used to trace relationships amongst academic journal citations. Now, bibliometrics are important in indexing research performance.

Bibliometric data have particular characteristics of which the user should be aware, and these are considered here.

Journal papers (publications, sources) report research work. Papers refer to or 'cite' earlier work relevant to the material being reported. New papers are cited in their turn. Papers that accumulate more citations are thought of as having greater 'impact', which is interpreted as significance or influence on their field. Citation counts are therefore recognised as a measure of impact, which can be used to index the excellence of the research from a particular group, institution or country.

The origins of citation analysis as a tool that could be applied to research performance can be traced to the mid-1950s, when Eugene Garfield proposed the concept of citation indexing and introduced the Science Citation Index, the Social Sciences Citation Index and the Arts & Humanities Citation Index, produced by the Institute of Scientific Information (currently the IP & Science business of Clarivate Analytics).<sup>14</sup>

We can count citations, but they are only 'indicators' of impact or quality – not metrics. Most impact indicators use average citation counts from groups of papers, because some individual papers may have unusual or misleading citation profiles. These outliers are diluted in larger samples.

---

### DATA SOURCE

The data we use come from the Clarivate Analytics *Web of Science*<sup>™</sup> databases which give access not only to journals but also to conference proceedings, books, patents, websites, and chemical structures, compounds and reactions. It has a unified structure that integrates all data and search terms together and therefore provides a level of comparability not found in other databases. It is widely acknowledged to be the world's leading source of citation and bibliometric data. The *Web of Science* focuses on research published in journals, conferences and books in science, medicine, arts, humanities and social sciences.

The *Web of Science* was originally created as an awareness and information retrieval tool but it has acquired an important primary use as a tool for research evaluation, using citation analysis and bibliometrics. Data coverage is both current and retrospective in the sciences, social sciences, arts and humanities, in some cases back to 1900. Within the research community this data source was previously referred to by the acronym 'ISI'.

---

<sup>14</sup>Garfield, E (1955) Citation Indexes for Science – New dimension in documentation through association of ideas. *Science*: 122, 108-111.

Unlike other databases, the *Web of Science* and underlying databases are selective, that is: the journals abstracted are selected using rigorous editorial and quality criteria. The authoritative, multidisciplinary content covers over 12,000 of the highest impact journals worldwide, including Open Access journals, and over 150,000 conference proceedings. The abstracted journals encompass the majority of significant, frequently cited scientific reports and, more importantly, an even greater proportion of the scientific research output which is cited. This selective process ensures that the citation counts remain relatively stable in given research fields and do not fluctuate unduly from year to year, which increases the usability of such data for performance evaluation.

Clarivate Analytics has extensive experience with databases on research inputs, activity and outputs and has developed innovative analytical approaches for benchmarking and interpreting international, national and institutional research impact.

## DATABASE CATEGORIES

The source data can be grouped in various classification systems. Most of these are based on groups of journals that have a relatively high cross-citation linkage and naturally cluster together. Custom classifications use subject maps in third-party data such as the OECD categories set out in the Frascati manual.

Clarivate Analytics frequently uses the broader field categories in the Essential Science Indicators system and the finer journal categories in the Web of Science. There are 22 fields in Essential Science Indicators and 254 fields in Web of Science. In either case, our bibliometric analyses draw on the full range of data available in the underlying database, so analyses in our reports will differ slightly from anything created 'on the fly' from data in the web interface.

The lists of journal categories in these systems are attached at the end of this document.

Most analyses start with an overall view across the data, then move to a view across broad categories and only then focus in at a finer level in the areas of greatest interest to policy, programme or organisational purpose.

## ASSIGNING PAPERS TO ADDRESSES

A paper is assigned to each country and each organisation whose address appears at least once for any author on that paper. One paper counts once and only once for each assignment, however many address variants occur for the country or organisation. No weighting is applied.

For example, a paper has five authors, thus:

Author	Organisation	Country		
Gurney, KA	Univ Leeds	UK	Counts for Univ Leeds	Counts for UK
Adams, J	Univ Leeds	UK	No gain for Univ Leeds	No gain for UK
Kochalko, D	Univ C San Diego	USA	Counts for UCSD	Counts for USA
Munshi, S	Gujarat Univ	India	Counts for Gujarat Univ	Counts for India
Pendlebury, D	Univ Oregon	USA	Counts for Univ Oregon	No gain for USA

So this one paper with five authors would be included once in the tallies for each of four universities and once in the tallies for each of three countries.

Work carried out within Clarivate Analytics, and research published elsewhere, indicates that fractional weighting based on the balance of authors by organisation and country makes little difference to the conclusions of an analysis at an aggregate level. Such fractional analysis can introduce unforeseen errors in the attempt to create a detailed but uncertain assignment. Partitioning credit would make a greater difference at a detailed, group level but the analysis can then be manually validated.

---

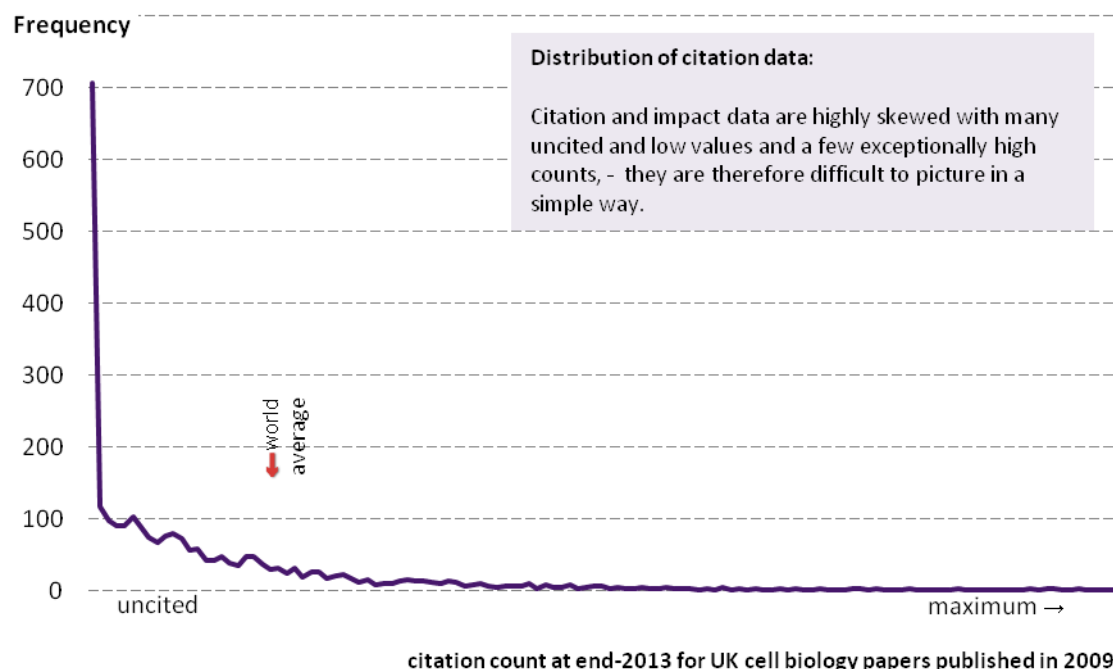
## CITATION COUNTS

A publication accumulates citation counts when it is referred to by more recent publications. Some papers get cited frequently and many get cited rarely or never, so the distribution of citations is highly skewed.

Why are many papers never cited? Certainly some papers remain uncited because their content is of little or no impact, but that is not the only reason. It might be because they have been published in a journal not read by researchers to whom the paper might be interesting. It might be that they represent important but 'negative' work reporting a blind alley to be avoided by others. The publication may be a commentary in an editorial, rather than a normal journal article and thus of general rather than research interest. Or it might be that the work is a 'sleeping beauty' that has yet to be recognised for its significance.

Other papers can be very highly cited: hundreds, even thousands of times. Again, there are multiple reasons for this. Most frequently cited work is being recognised for its innovative significance and impact on the research field of which it speaks. Impact here is a good reflection of quality: it is an indicator of excellence. But there are other papers which are frequently cited because their significance is slightly different: they describe key methodology; they are a thoughtful and wide-ranging review of a field; or they represent contentious views which others seek to refute.

Citation analysis cannot make value judgments about why an article is uncited nor about why it is highly cited. The analysis can only report the citation impact that the publication has achieved. We normally assume, based on many other studies linking bibliometric and peer judgments, that high citation counts correlate on average with the quality of the research.



The figure shows the skewed distribution of more or less frequently cited papers from a sample of UK authored publications in cell biology. The skew in the distribution varies from field to field. It is to compensate for such factors that actual citation counts must be normalised, or rebased, against a world baseline.

We do not seek to account separately for the effect of self-citation. If the citation count is significantly affected by self-citation then the paper is likely to have been infrequently cited. This is therefore only of consequence for low impact activity. Studies show that for large samples at national and organisational level the effect of self-citation has little or no effect on the analytical outcomes and would not alter interpretation of the results.

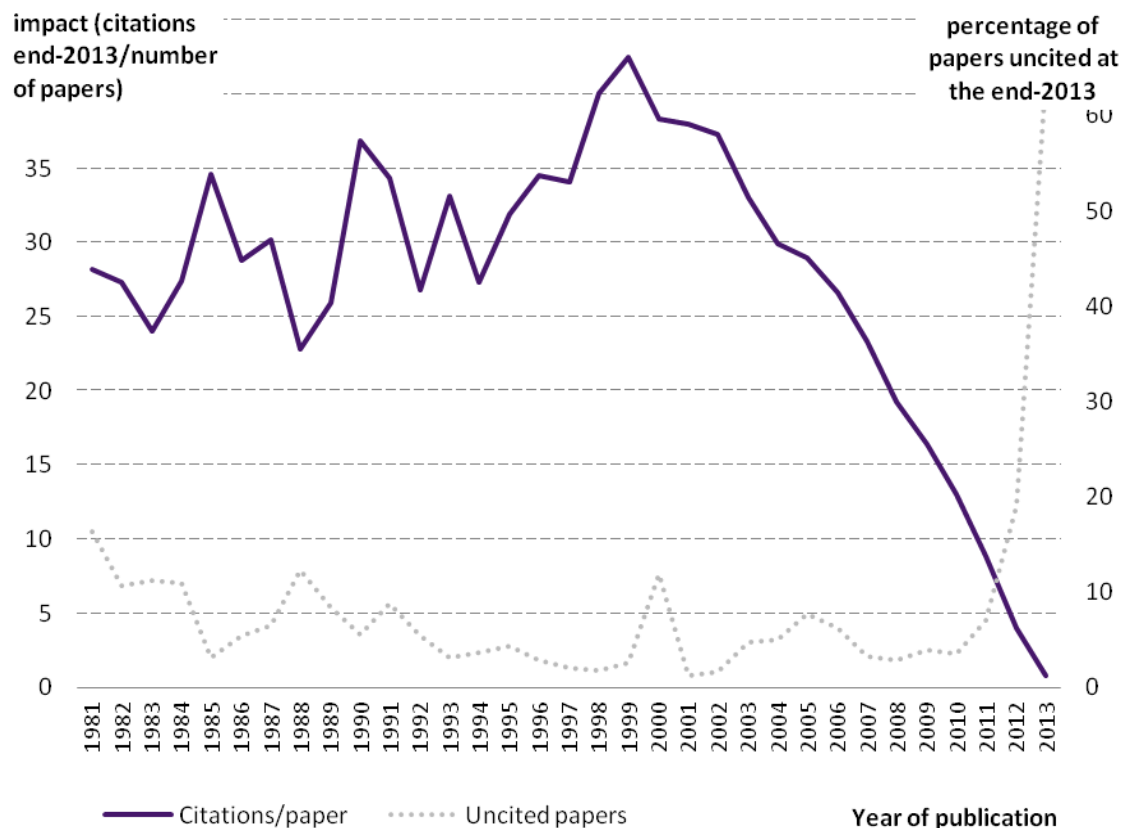
## TIME FACTORS

Citations accumulate over time. Older papers therefore have, on average, more citations than more recent work. The graph below shows the pattern of citation accumulation for a set of 33 journals in the journal category **Materials Science, Biomaterials**. Papers less than eight years old are, on average, still accumulating additional citations. The citation count goes on to reach a plateau for older sources.

The graph shows that the percentage of papers that have never been cited drops over about five years. Beyond five years, between 5% and 10% or more of papers remain uncited.

Account must be taken of these time factors in comparing current research with historical patterns. For these reasons, it is sometimes more appropriate to use a fixed five-year window of papers and citations to compare two periods than to look at the longer term profile of citations and of uncitedness for a recent year and an historical year.





## DISCIPLINE FACTORS

Citation rates vary between disciplines and fields. For the UK science base as a whole, ten years produces a general plateau beyond which few additional citations would be expected. On the whole, citations accumulate more rapidly and plateau at a higher level in biological sciences than physical sciences, and natural sciences generally cite at a higher rate than social sciences.

Papers are assigned to disciplines (journal categories or research fields) by Clarivate Analytics, bringing cognate research areas together. The journal category classification scheme has been recently revised and updated. Before 2007, journals were assigned to the older, well established Current Contents categories which were informed by extensive work by Thomson and with the research community since the early 1960s. This scheme has been superseded by the 252 *Web of Science*<sup>SM</sup> journal categories which allow for greater disaggregation for the growing volume of research which is published and abstracted.

Papers are allocated according to the journal in which the paper is published. Some journals may be considered to be part of the publication record for more than one research field. As the example below illustrates, the journal *Acta Biomaterialia* is assigned to two journal categories: **Materials Science, Biomaterials** and **Engineering, Biomedical**.

Very few papers are not assigned to any research field and as such will not be included in specific analyses using Normalised citation impact data. The journals included in the Clarivate Analytics databases and how they are selected are detailed here <http://scientific.thomsonreuters.com/mjl/>.

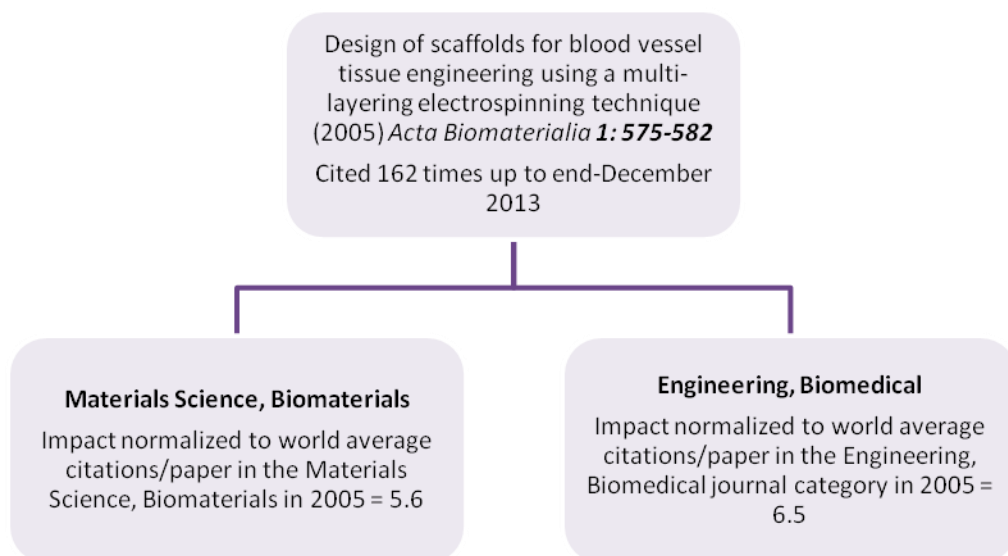
Some journals with a very diverse content, including the prestigious journals *Nature* and *Science* were classified as **Multidisciplinary** in databases created prior to 2007. The papers from these **Multidisciplinary** journals are now re-assigned to more specific research fields using an algorithm based on the research area(s) of the references cited by the article.

## NORMALIZED CITATION IMPACT

Because citations accumulate over time at a rate that is dependent upon the field of research, all analyses must take both field and year into account. In other words, because the absolute citation count for a specific article is influenced by its field and by the year it was published, we can only make comparisons of indexed data after normalising with reference to these two variables.

We only use citation counts for reviews and articles in calculations of impact, because document type influences the citation count. For example, a review will often be cited more frequently than an article in the same field, but editorials and meeting abstracts are rarely cited and citation rates for conference proceedings are extremely variable. The most common normalisation factors are the average citations per paper for (1) the year and (2) either the field or the journal in which the paper was published. This normalisation is also referred to as ‘rebasings’ the citation count.

Impact is therefore most commonly analysed in terms of ‘Normalised impact’, or NCI. The following schematic illustrates how the Normalised citation impact is calculated at paper level and journal category level.



This article in the journal *Acta Biomaterialia* is assigned to two journal categories: **Materials Science, Biomaterials** and **Engineering, Biomedical**. The world average baselines for, as an example, **Materials science, Biomaterials** are calculated by summing the citations to all the articles and reviews published worldwide in the journal *Acta Biomaterialia* and the other 32 journals assigned to this category for each year, and dividing this by the total number of articles and reviews published in the journal category. This gives the category-specific Normalised citation impact (in the above example the category-specific  $NCI_F$  for **Materials Science, Biomaterials** is 5.6 and the category-

specific NCI<sub>F</sub> for **Engineering, Biomedical** is higher at 6.5). Most papers (nearly two-thirds) are assigned to a single journal category whilst a minority are assigned to more than 5.

Citation data provided by Clarivate Analytics are assigned on an annual census date referred to as the Article Time Period. For the majority of publications the Article Time Period is the same as the year of publication, but for a few publications (especially those published at the end of the calendar year in less main-stream journals) the Article Time Period may vary from the actual year of publication.

World average impact data are sourced from the Clarivate Analytics National Science Indicators baseline data for 2013.

---

### MEAN NORMALIZED CITATION IMPACT

Research performance has historically been indexed by using average citation impact, usually compared to a world average that accounts for time and discipline. As noted, however, the distribution of citations amongst papers is highly skewed because many papers are never cited while a few papers accumulate very large citation counts. That means that an average may be misleading if assumptions are made about the distribution of the underlying data.

In fact, almost all research activity metrics are skewed: for research income, PhD numbers and publications there are many low activity values and a few exceptionally high values. In reality, therefore, the skewed distribution means that average impact tends to be greater than and often significantly different from either the median or mode in the distribution. This should be borne in mind when reviewing analytical outcomes.

The average (Normalised) citation impact can be calculated at an individual paper level where it can be associated with more than one journal category. It can also be calculated for a set of papers at any level from a single country to an individual researcher's output. In the example above, the average citation impact of the *Acta Biomaterialia* paper can be expressed as  $((5.6 + 6.5)/2) = 6.1$ .

---

### IMPACT PROFILES®

We have developed a bibliometric methodology<sup>15</sup> that shows the proportion of papers that are uncited and the proportion that lie in each of eight categories of relative citation rates, Normalised (rebased) to world average. An Impact Profile® enables an examination and analysis of the strengths and weaknesses of published outputs relative to world average and relative to a reference profile. This provides much more information about the basis and structure of research performance than conventionally reported averages in citation indices.

Papers which are “highly-cited” are often defined in our reports as those with an average citation impact (NCI<sub>F</sub>) greater than or equal to 4.0, i.e. those papers which have received greater than or equal to four times the world average number of citations for papers in that subject published in that year. This differs from Clarivate Analytics database of global highly-cited papers, which are the top 1% most frequently cited for their field and year. The top percentile is a powerful indicator of leading performance but is too stringent a threshold for most management analyses.

The proportion of uncited papers in a dataset can be compared to the benchmark for the UK, the USA or any other country. Overall, in a typical ten-year sample, around one-quarter of papers have not

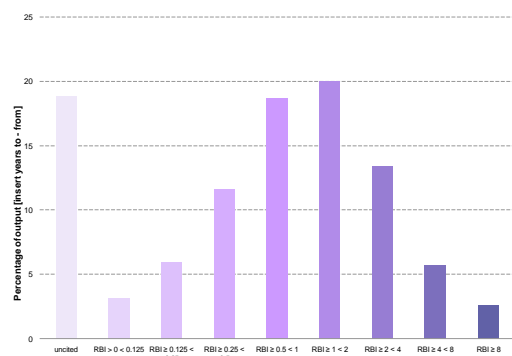
---

<sup>15</sup>Adams J, Gurney K & Marshall S (2007) Profiling citation impact: A new methodology. *Scientometrics* 72: 325-344.

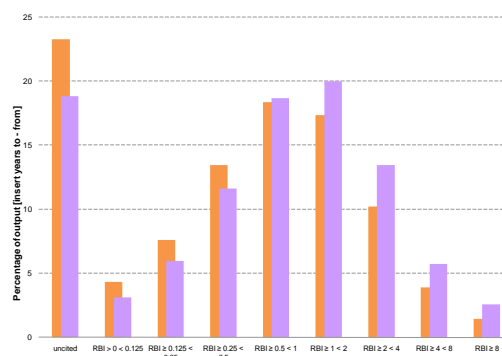
been cited within the 10-year period; the majority of these are, of course, those that are most recently published.

The Impact Profile® histogram can be presented in a number of ways which are illustrated below.

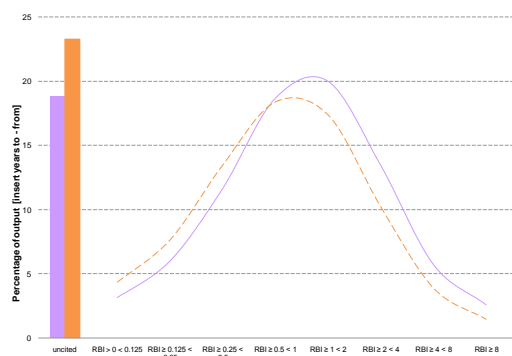
A



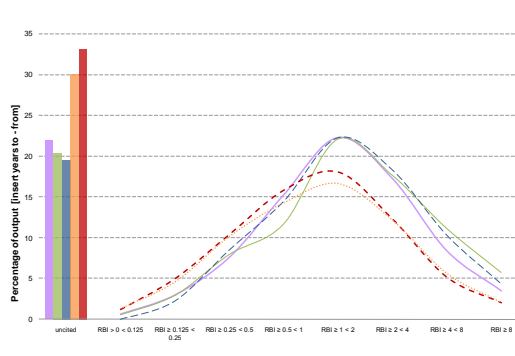
B



C



D



A: is used to represent the total output of an individual country, institution or researcher with no benchmark data. Visually it highlights the numbers of uncited papers (weaknesses) and highly cited papers (strengths).

B & C: are used to represent the total output of an individual country, institution or researcher (client) against an appropriate benchmark dataset (benchmark). The data are displayed as either histograms (B) or a combination of histogram and profile (C). Version C prevents the 'travel' which occurs in histograms where the eye is drawn to the data most offset to the right, but can be less easy to interpret as categorical data.

D: illustrates the complexity of data which can be displayed using an Impact Profile®. These data show research output in defined journal categories against appropriate benchmarks: client, research field X; client, research field Y; client, research field Z; benchmark, research field X+Y; benchmark, research field, Z.

Impact Profiles® enable an examination and analysis of the balance of published outputs relative to world average and relative to a reference profile. This provides much more information about the basis and structure of research performance than conventionally reported averages in citation indices.

An Impact Profile® shows what proportion of papers are uncited and what proportion are in each of eight categories of relative citation rates, Normalised to world average (which becomes 1.0 in this graph). Normalised citation rates above 1.0 indicate papers cited more often than world average for the field in which that journal is categorised and in their year of publication.

Attention should be paid to:

- The proportion of uncited papers on the left of the chart
- The proportion of cited papers either side of world average (1.0)
- The location of the most common (modal) group near the centre
- The proportion of papers in the most highly-cited categories to the right, ( $\geq 4 \times$  world,  $\geq 8 \times$  world).

---

#### WHAT ARE UNCITED PAPERS?

It may be a surprise that some journal papers are never subsequently cited after publication, even by their authors. This accounts for about half the total global output for a typical, recent 10-year period. We cannot tell why papers are not cited. It is likely that a significant proportion of papers remain uncited because they are reporting negative results which are an essential matter of record in their field but make the content less likely to be referenced in other papers. Inevitably, other papers are uncited because their content is trivial or marginal to the mainstream. However, it should not be assumed that this is the case for all such papers.

There is variation in non-citation between countries and between fields. For example, relatively more engineering papers tend to remain uncited than papers in other sciences, indicative of a disciplinary factor but not a quality factor. While there is also an obvious increase in the likelihood of citation over time, most papers that are going to be cited will be cited within a few years of publication.

---

#### WHAT IS THE THRESHOLD FOR 'HIGHLY CITED'?

Clarivate Analytics has traditionally used the term 'Highly Cited Paper' to refer to the world's 1% of most frequently cited papers, taking into account year of publication and field. In rough terms, UK papers cited more than eight times as often as relevant world average would fall into the Clarivate Highly Cited category. About 1-2% of papers (all papers, cited or uncited) typically pass this hurdle. Such a threshold certainly delimits exceptional papers for international comparisons but, in practice, is an onerous marker for more general management purposes.

After reviewing the outcomes of a number of analyses, we have chosen a more relaxed definition for our descriptive and analytical work. We deem papers that are in the world's top 10% of most frequently cited papers, taking into account year of publication and field, to be relatively highly-cited for national comparisons.