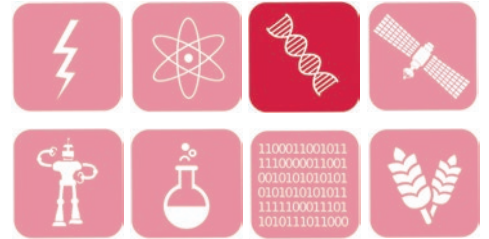




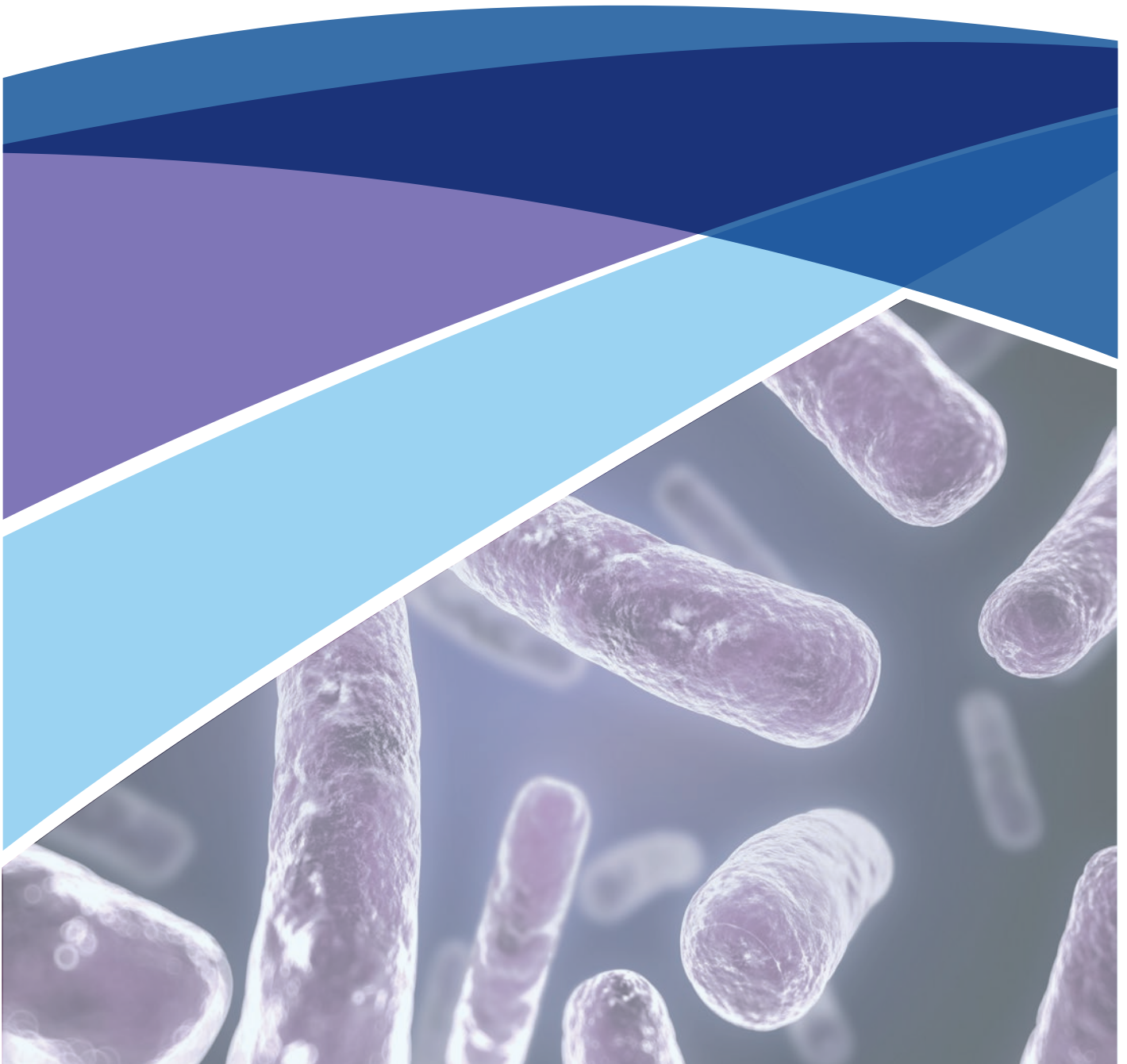
Intellectual  
Property  
Office



# Eight Great Technologies

## Regenerative Medicine

A patent overview



## #8Great

This report was prepared by the  
UK Intellectual Property Office Informatics Team  
June 2014

e-mail: [informatics@ipo.gov.uk](mailto:informatics@ipo.gov.uk)

© Intellectual Property Office 2014  
Intellectual Property Office  
Concept House  
Cardiff Road  
Newport  
NP10 8QQ  
United Kingdom

[www.ipo.gov.uk/informatics](http://www.ipo.gov.uk/informatics)



# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Worldwide patent analysis</b>	<b>3</b>
2.1	Overview	3
2.2	Top applicants	10
2.3	Collaboration	12
2.4	Technology breakdown	15
<b>3</b>	<b>The UK landscape</b>	<b>17</b>
3.1	Top UK applicants	17
3.2	UK inventor mobility	18
3.3	How active is the UK?	19
<b>4</b>	<b>Patent landscape map analysis</b>	<b>21</b>
<b>5</b>	<b>Conclusions</b>	<b>23</b>
Appendix A	Interpretation notes	24
Appendix B	Relative Specialisation Index	26
Appendix C	Patent landscape maps	27

# 1 Introduction

The UK Government has identified 'eight great technologies' which will propel the UK to future growth. These are:

- the big data revolution and energy-efficient computing;
- satellites and commercial applications of space;
- robotics and autonomous systems;
- life sciences, genomics and synthetic biology;
- regenerative medicine;
- agri-science;
- advanced materials and nanotechnology;
- energy and its storage.

Patent data can give a valuable insight into innovative activity, to the extent that it has been codified in patent applications, and the IPO Informatics team is producing a series of patent landscape reports looking at each of these technology spaces and the current level of UK patenting on the world stage. As an aid to help people understand the eight great technologies and to consider the direction of future funding, the IPO is offering a comprehensive overview of what is already patented in the each of these technologies and in which direction the technology is developing.

This report analyses the worldwide patent landscape for regenerative medicine. Regenerative medicine involves replacing or restoring cells, tissues, or organs in the body. This includes transplanting new cells, tissues, or organs into the body, stimulating the body's own self-repair mechanisms, and developing new materials for structural repairs. The dataset used for analysis was extracted from worldwide patent databases following detailed discussion and consultation with patent examiners from the Intellectual Property Office who are experts in the field and who, on a day-to-day basis, search, examine and grant patent applications relating to medical technologies.

This report is based on analysis of published patent application data and not granted patent data. Data for published patent applications gives more information about technological activity than the figures for granted patents because a number of factors determine whether an application ever proceeds to grant. These include the inherent lag in patent processing at national IP offices worldwide and the patenting strategies of applicants who may file more applications than they ever intend to pursue.

## 2 Worldwide patent analysis

### 2.1 Overview

Table 1 gives a summary of the extracted and cleaned dataset used for this analysis of regenerative medicine. All of the analysis undertaken in this report was performed on this dataset or a subset of this dataset. The worldwide dataset for regenerative medicine published between 2003 and 2013 contains 70,500 published patents equating to almost 10,000 patent families. Publications may be at the application or grant stage, so are not necessarily granted patents. A patent family is one or more published patents originating from a single original (priority) application. Analysis by patent family more accurately reflects the number of inventions present because generally there is one invention per patent family, whereas analysis by raw number of patent publications inevitably involves double counting because one patent family may contain dozens of patent publications if the applicant files for the same invention in more than one country. Hence analysis by patent family gives more accurate results regarding the level of inventive activity taking place.

**Table 1: Summary of worldwide patent dataset for regenerative medicine**

<b>Number of patent families</b>	9912		
<b>Number of patent publications</b>	70,500		
<b>Publication year range</b>	2004-2013		
<b>Peak publication year</b>	2012		
<b>Top applicant</b>	University of California (USA)		
<b>Field choices</b>	<b>Field name</b>	<b>Number of entries</b>	<b>Coverage</b>
<b>People</b>	Inventors	26,647	99%
<b>Applicants</b>	Patent assignees	4146	92%
<b>Countries</b>	Priority countries	43	100%
<b>Technology</b>	IPC sub-group	3950	100%

Figure 1 shows the total number of published patents by publication year (top) and the total number of patent families by priority year (bottom – considered to be the best indication of when the original invention took place). Figure 1 suggests that patenting in regenerative medicine has grown slowly over the period and flattened most recently, although peak years by publication and priority are in 2012 and 2010 respectively. The patent family chart in red does not show any patents filed after 2011 because a patent is normally published 18 months after the priority date or the filing (application) date, whichever is earlier. Hence, the 2012 and 2013 data is incomplete and has been ignored.

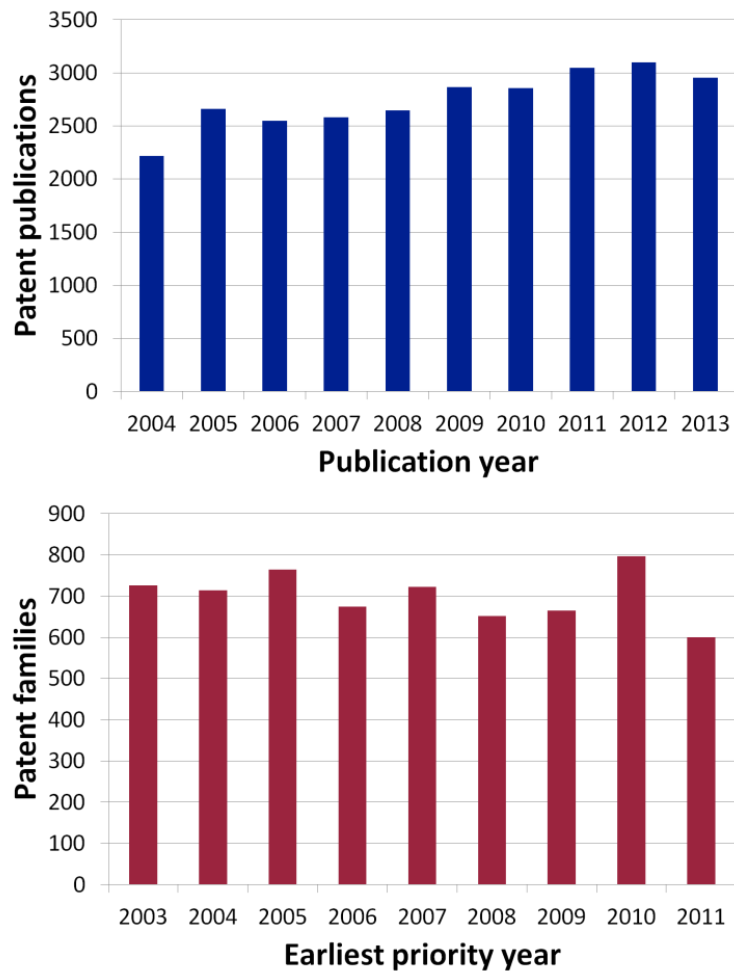
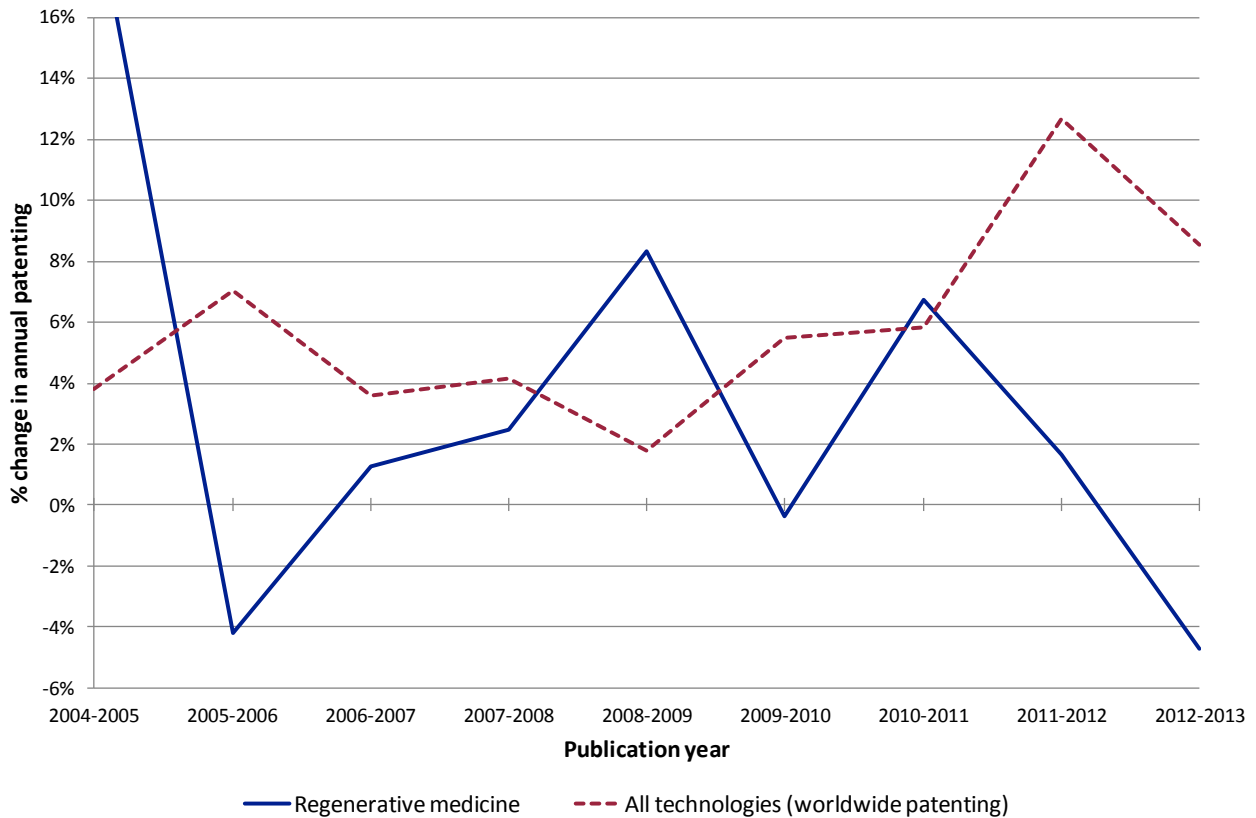


Figure 1: Patent publications by publication year (top) and patent families by priority year (bottom)

General patenting levels globally have continued to grow at an ever-increasing rate. Figure 2 tries to address this issue by normalising the data shown in Figure 1 and presenting the annual increase in the size of the worldwide patent databases across all technologies against the year-on-year change in the size of the regenerative medicine dataset. For example, from 2012 to 2013 worldwide patenting across all areas of technology increased by 8.6% and this can be compared to a 4.7% decrease in regenerative medicine patenting over the same time period. The data suggest that that patenting in regenerative medicine is being overtaken by patenting in general.



**Figure 2: Year-on-year change in regenerative medicine patenting compared to worldwide patenting across all technologies**

Figure 3 shows the priority country distribution across the dataset with almost two thirds of regenerative medicine patents first filed in the USA. 3% of regenerative medicine-related patents are first filed in the UK. Traditionally priority country analysis has been a good indicator of where the invention is actually taking place because many applicants will file patent applications first in the country in which they reside<sup>1</sup>, but in recent years drawing firm conclusions from this data is harder because there may be other strategic reasons for an applicant choosing the country of first filing (e.g. tax treatment).

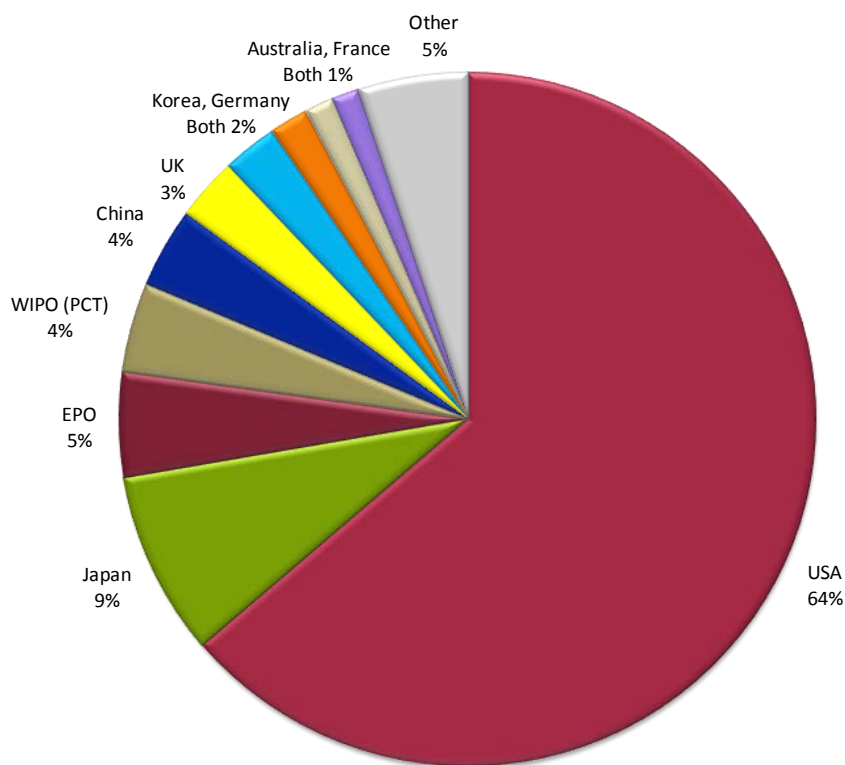


Figure 3: Priority country distribution

It is interesting to compare the priority country distribution shown in Figure 3 and the applicant country distribution shown in Figure 4. The applicant country distribution shows a much greater diversity than the priority country distribution. 50% of patent families have an applicant living in the USA, yet 64% of patent families have a USA priority, and many of the remaining segments of the priority distribution are consequently enlarged (including the “other” countries). This demonstrates the significance of patenting in the USA, and the significance of the USA market, for applicants based elsewhere. Note that EPO and WIPO<sup>2</sup> may exist as priority countries but not as applicant countries.

<sup>1</sup>In some countries this is/was a requirement (e.g. in the UK this was a requirement until 2005).

<sup>2</sup>Alternative filing routes to single national patents, as outlined in Appendix A.3.



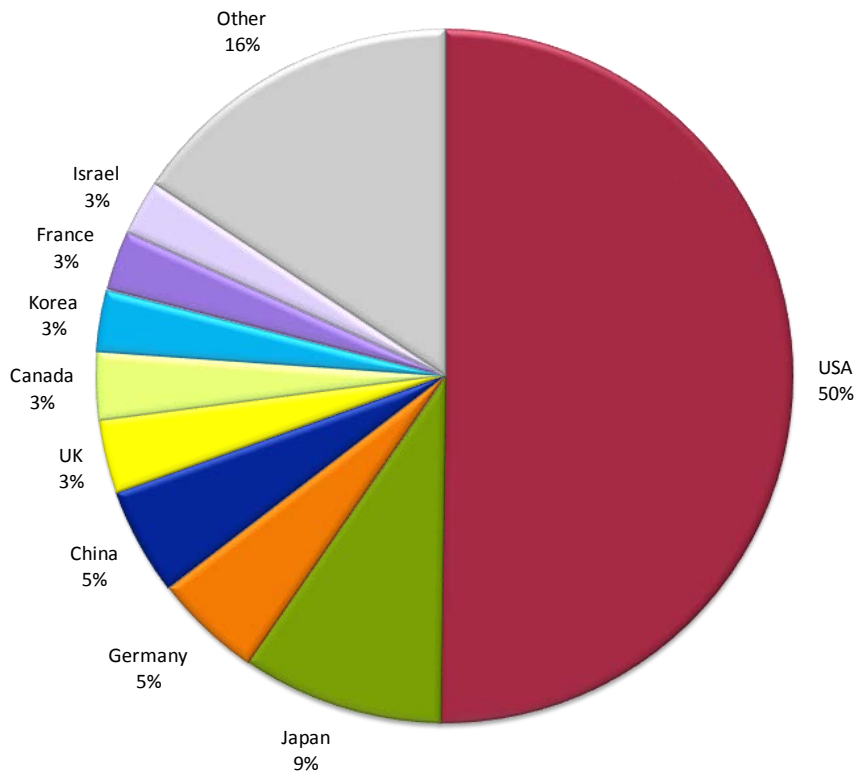


Figure 4: Applicant country distribution

It is well known that there is a greater propensity to patent in certain countries than others, and the trends shown in Figure 4 may change if the figures are corrected for this difference in behaviour. A Relative Specialisation Index (RSI)<sup>3</sup> for each applicant country has been calculated to give an indication of the level of invention in regenerative medicine technologies for each country compared to the overall level of invention in that country, and is shown in Figure 5.

The RSI shown in Figure 5 indicates that China, Korea and Germany in particular show a low specialisation in regenerative medicine. Israel, Australia, Canada and the USA show high specialisations. Israel, Australia, and Canada in particular have small shares of applicants in absolute terms but this picture highlights that they have a particular focus in regenerative medicine. The USA, on the other hand, has both a large share of applicants and a significant specialisation and is therefore quite dominant. The UK sits in the middle ground with some positive specialisation in regenerative medicine compared to other fields of patenting.

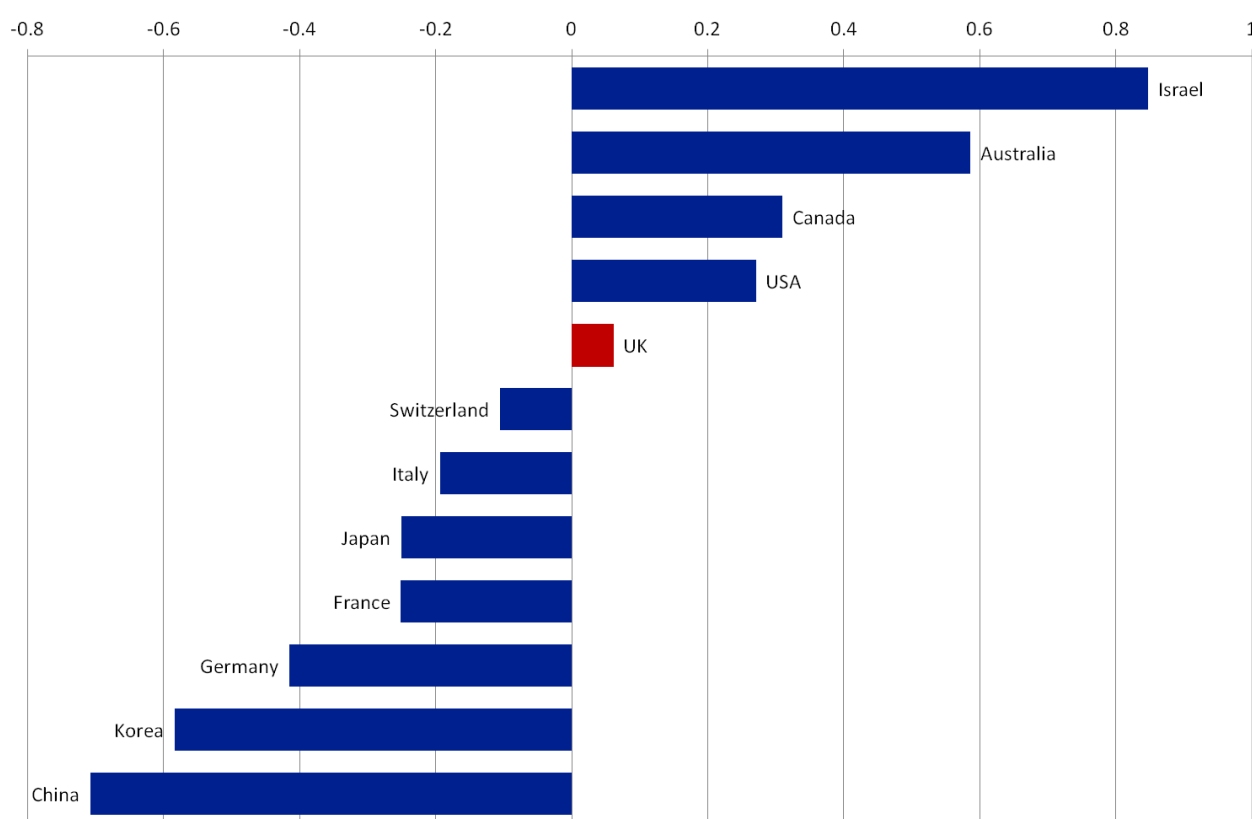


Figure 5: Relative Specialisation Index (RSI) by applicant country

<sup>3</sup> See Appendix B for full details of how the Relative Specialisation Index is calculated.

Figure 6 shows the countries in which applicants in the field of regenerative medicine are interested in seeking patent protection, with the strength of colour reflecting the quantity of published patents in each jurisdiction. Patents filed via the EPO [EPO] and WIPO (PCT) [WIPO] routes are also shown. Strong coverage is found in the USA, consistent with Figure 3 and Figure 4, and through the EPO and WIPO. Patent families with publications in the USA and through the EPO and WIPO have applicants from a wide range of other countries, however, indicating that there is a strong international dimension in regenerative medicine, with a significant degree of cross-border collaboration and knowledge flow. A high level of coverage is found in Canada and Australia, consistent with the high specialisation seen in those countries, but also in China, despite the low specialisation of Chinese residents. This indicates a significant interest by foreign applicants in protecting regenerative medicine inventions in China.

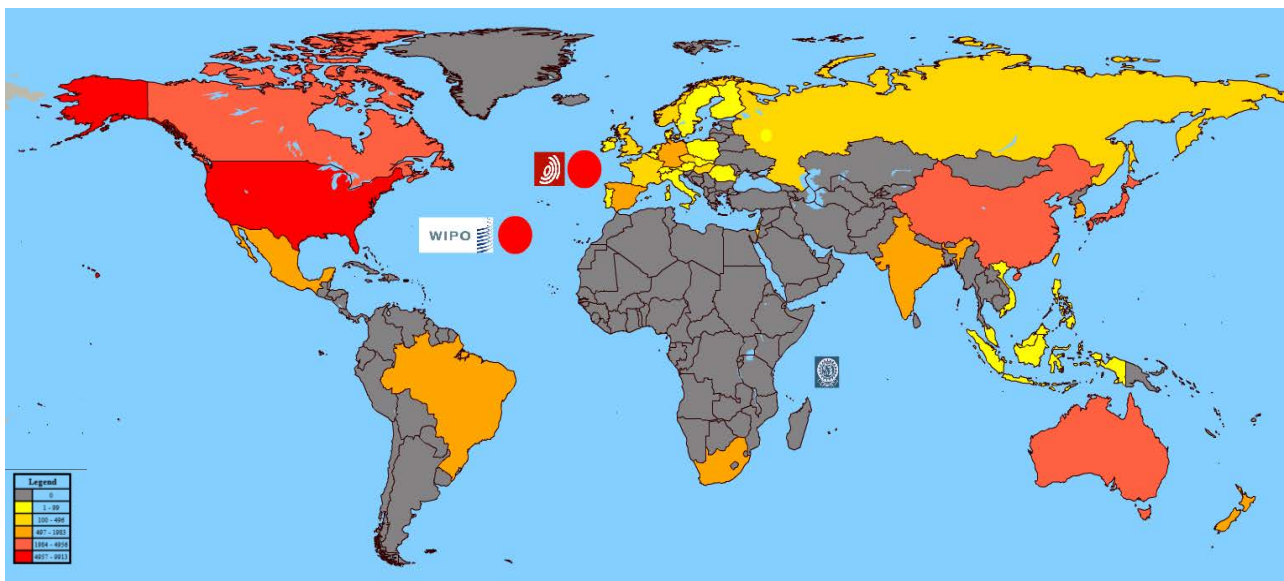


Figure 6: Patent coverage (publication country coverage)

## 2.2 Top applicants

Patent applicant names within the dataset were cleaned to remove duplicate entries arising from spelling errors, initialisation, international variation and equivalence<sup>4</sup>. Figure 7 shows the top 20 applicants in the regenerative medicine dataset.

As would be expected from the applicant distribution data already seen, many of the leading applicants are based in the USA. However, what is most striking is the number of academic institutions in this list, indicating that regenerative medicine is a highly research-based field, and hence an emerging technology with large further potential for growth as businesses enter the field and accumulate large patent portfolios. 2776 patent families out the total of 9913 (28%) contain a university applicant.

The University of California has a large lead over any other applicants.

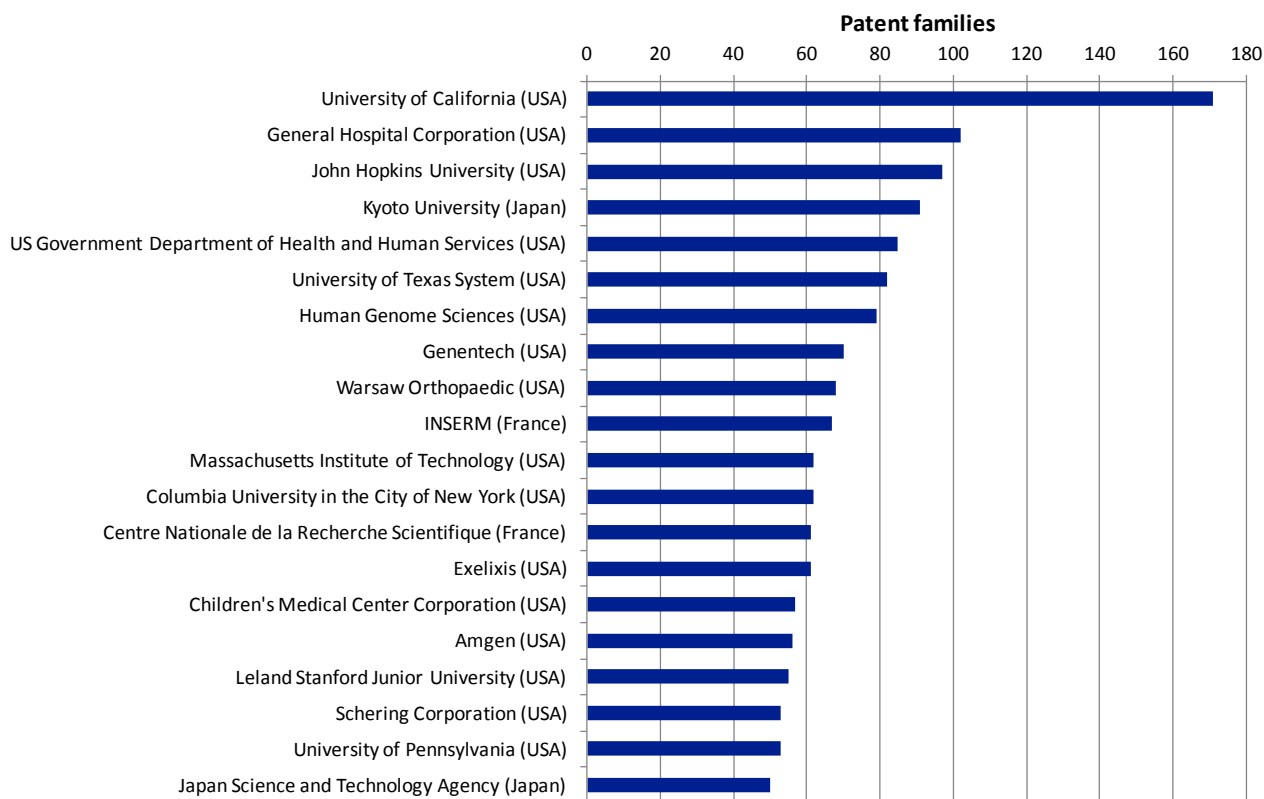


Figure 7: Top applicants

<sup>4</sup> See Appendix 0 for further details

Figure 8 is a bubble map showing a timeline for the top 20 applicants and shows the filing activity of these applicants in the last 10 years. It shows that most of the top applicants have been involved in regenerative medicine patenting throughout the period albeit in varying degrees. For example, the University of California and Kyoto University show large growth in recent years, whereas Amgen, Exelixis, Human Genome Sciences, and Schering have filed no applications in the last two years. A turnover of applicants is another indicator of an emerging technology as there are not yet any incumbents in the field.

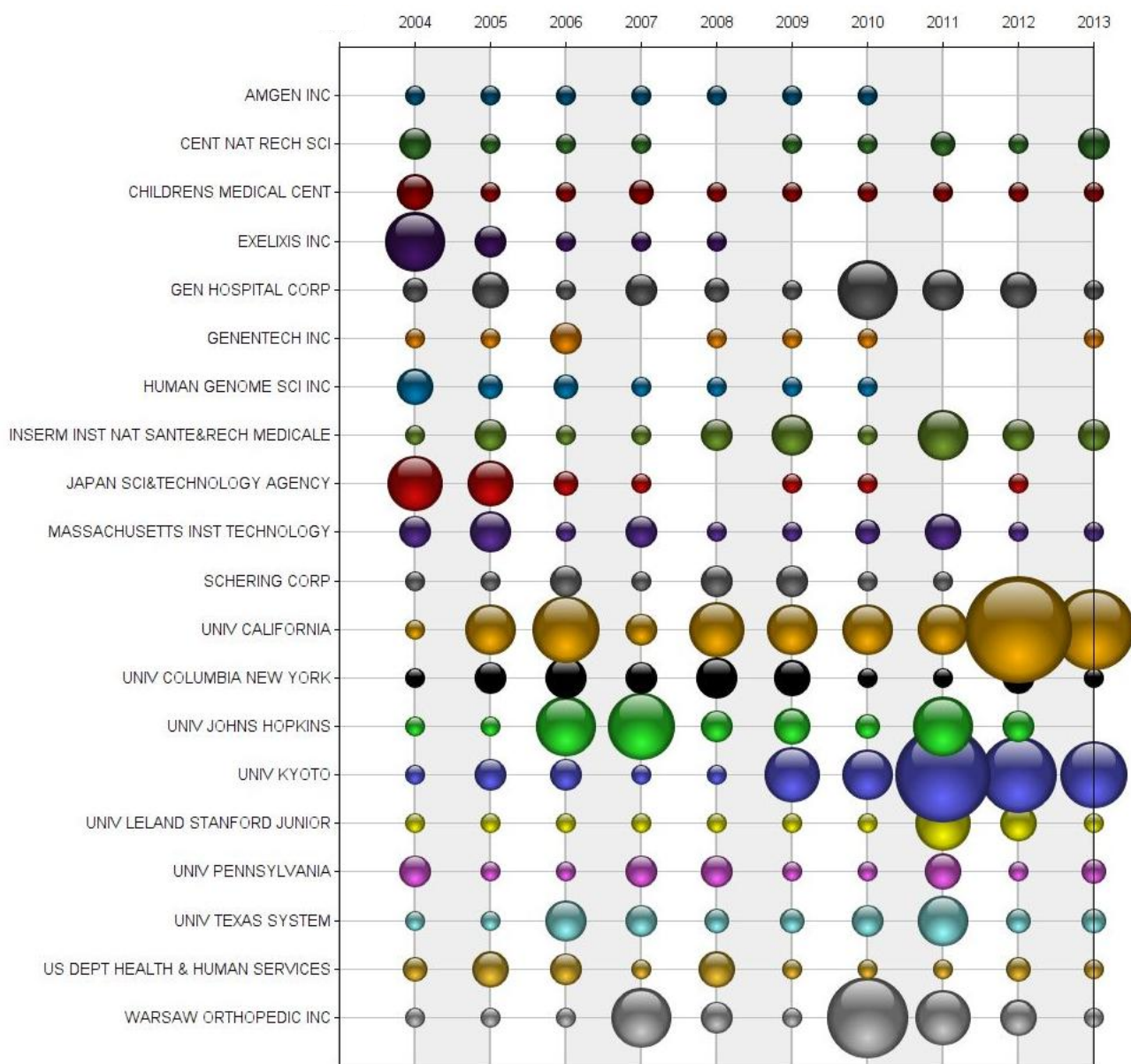


Figure 8: Applicant timeline of patent families by priority year

## 2.3 Collaboration

Figure 9 is a collaboration map showing all collaborations between the top five applicants in the dataset (the top five shown in Figure 7) and their collaborators. Each dot on the collaboration map represents a patent family and two applicants are linked together if they are named as joint applicants on a patent application. A collaboration map is a good indicator of technology transfer.

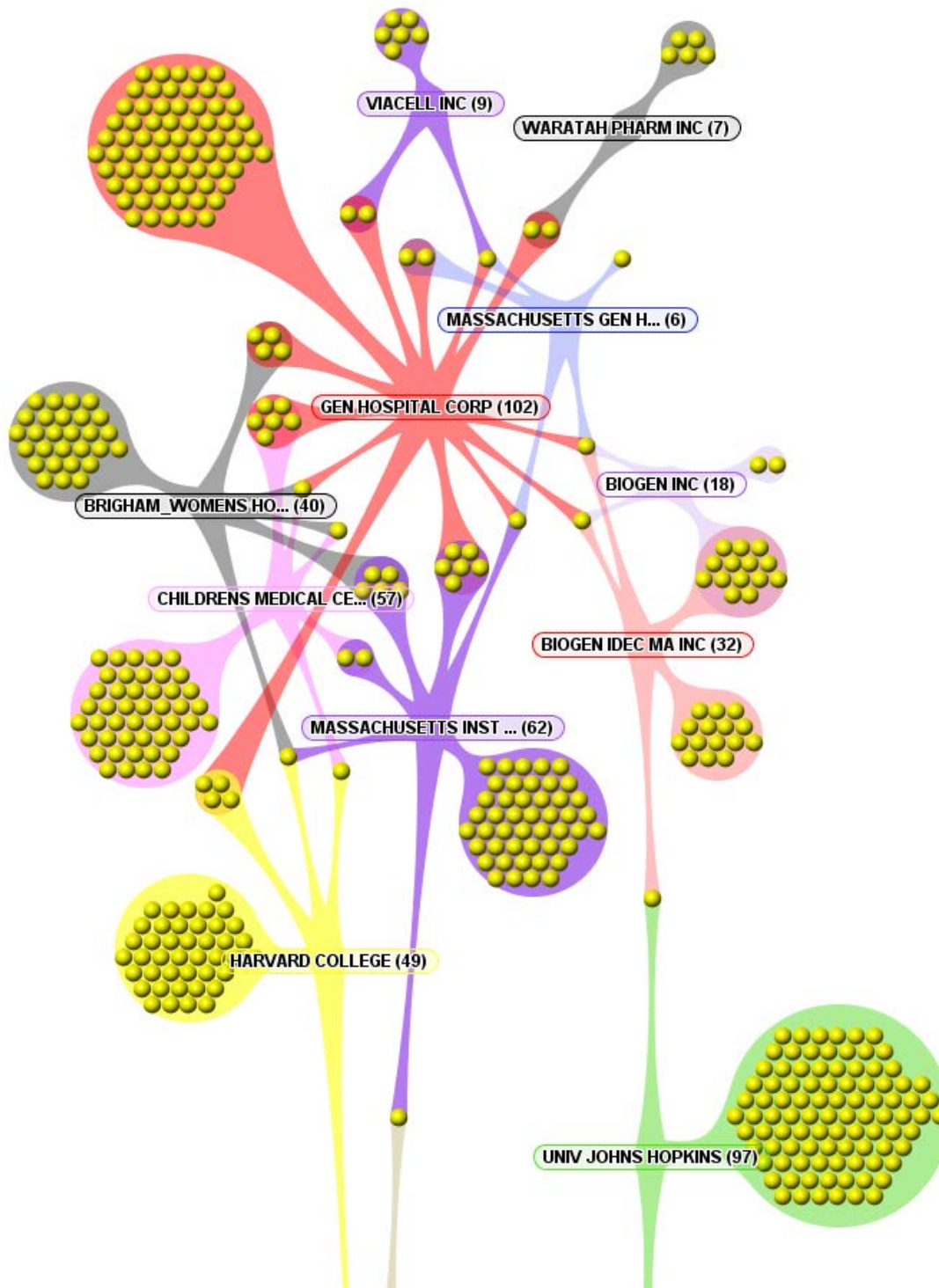


Figure 9 : Collaboration map showing all collaborations between the top 5 applicants and their collaborators



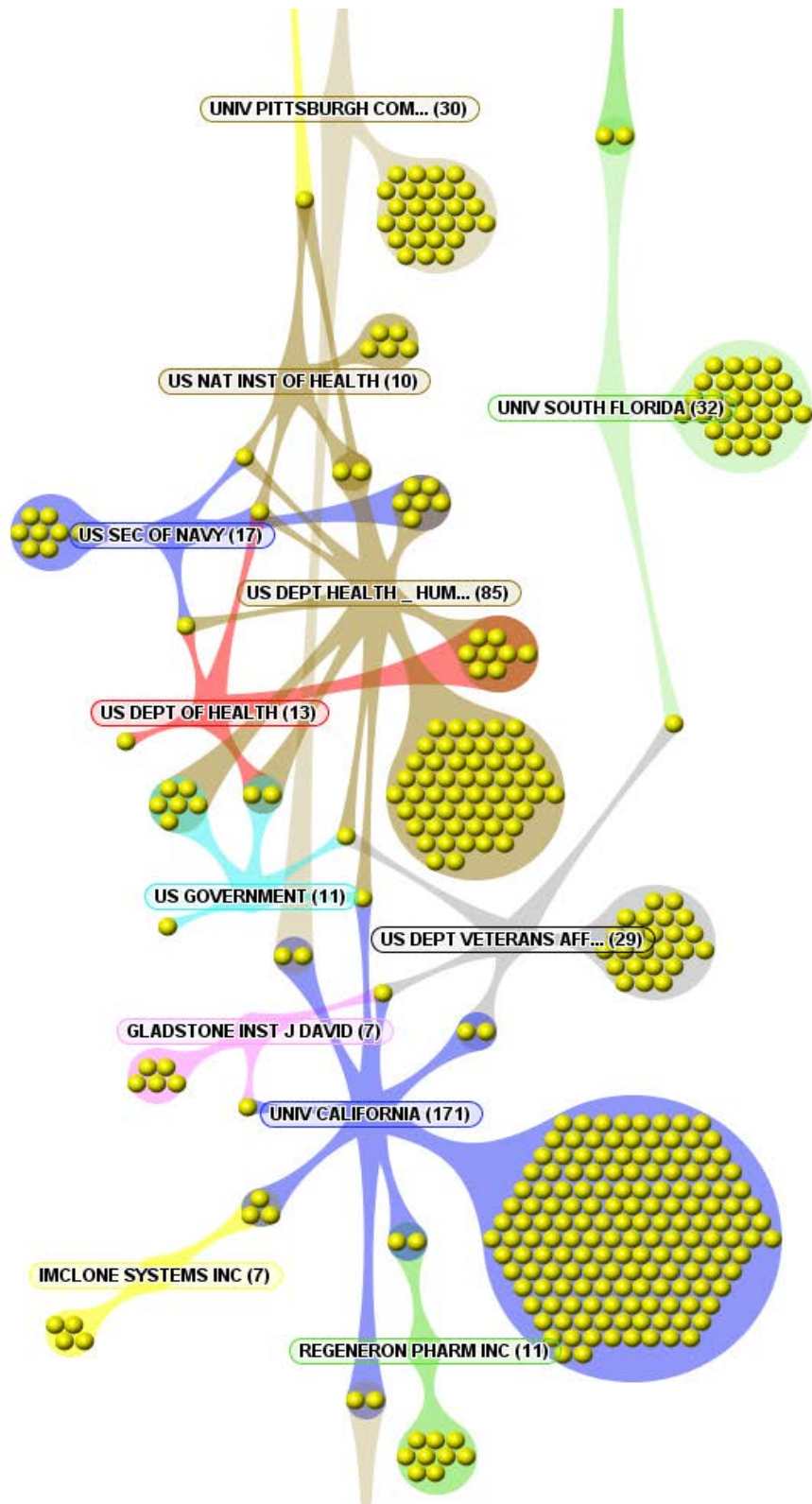


Figure 9 (cont.): Collaboration map showing all collaborations between the top 5 applicants and their collaborators

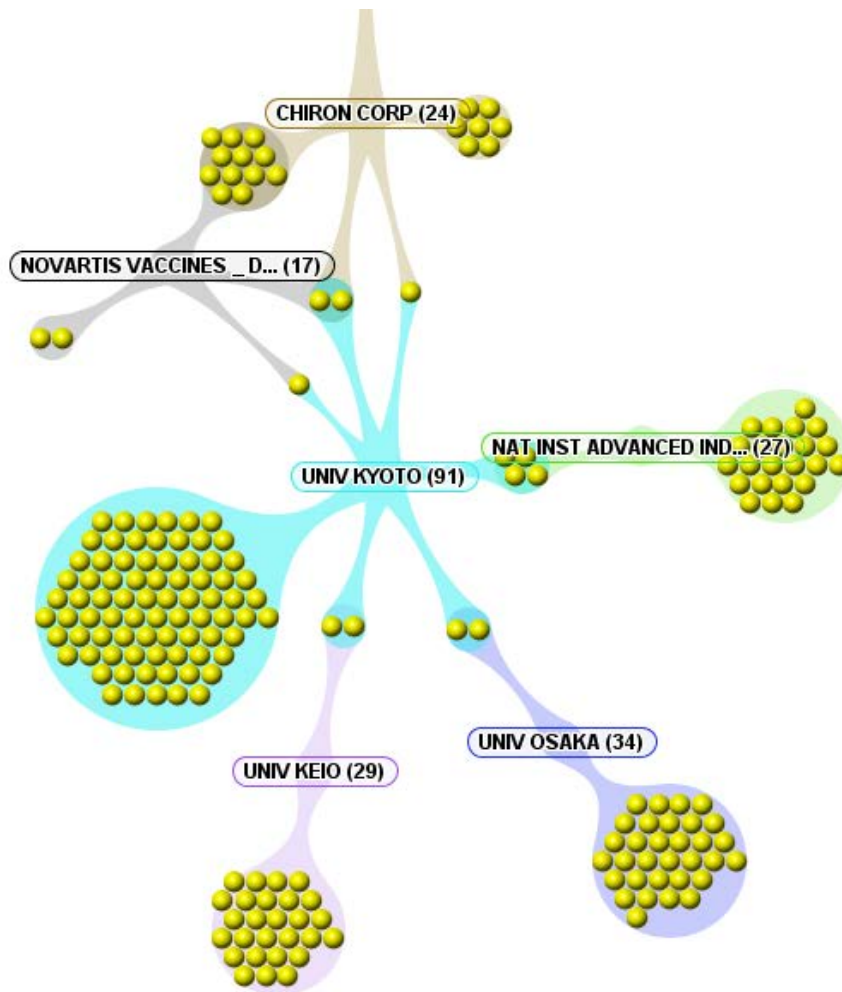


Figure 9 (cont.): Collaboration map showing all collaborations between the top 5 applicants and their collaborators

Figure 9 reveals a remarkable amount of collaboration and knowledge flow within the regenerative medicine field. Of the top five applicants, the University of California and the US Department for Health and Human Services are the only ones who collaborate directly, but it can be seen that all the names in this collaboration map are connected to each other at least indirectly, and there are no isolated clusters. Collaboration is seen between academic institutions, but also between academic institutions, hospitals, and businesses, and between the USA and Japan.

Figure 10 reveals a similar picture of large amounts of collaboration amongst UK applicants, despite the relatively small dataset size. In this case, however, there are distinct clusters, centred around the Glaxo Group (and its acquisitions) with Cambridge University spin-out companies, Imperial Innovation, Reneuron, the Medical Research Council, University College London, and the University of Edinburgh. There are several collaborations between universities, and internationally, particularly with the USA.



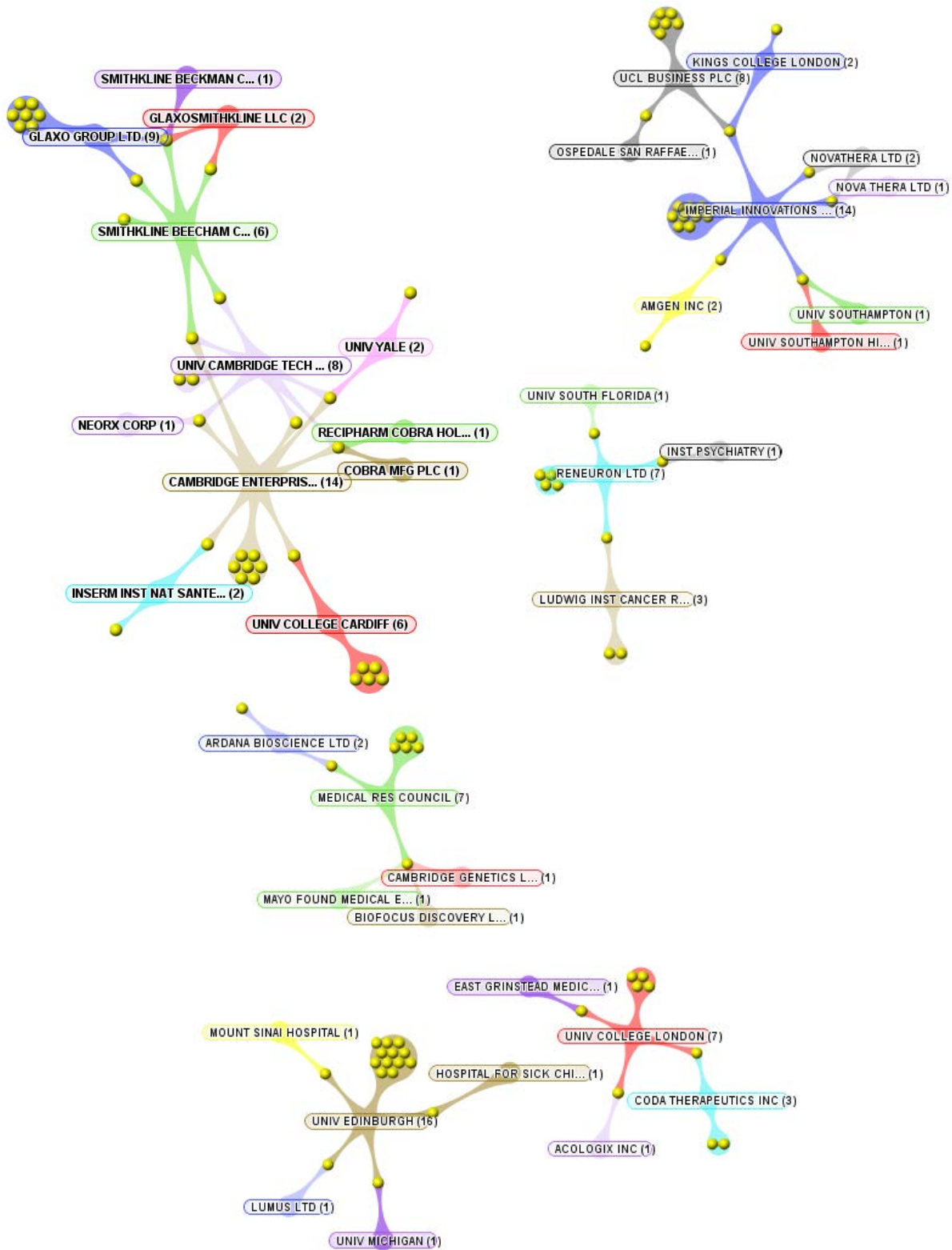


Figure 10 : Collaboration map showing all collaborations between the top 10 UK applicants and their collaborators

## 2.4 Technology breakdown

Figure 11 shows the top IPC sub-groups.

Table 2 lists the description of each of these sub-groups.

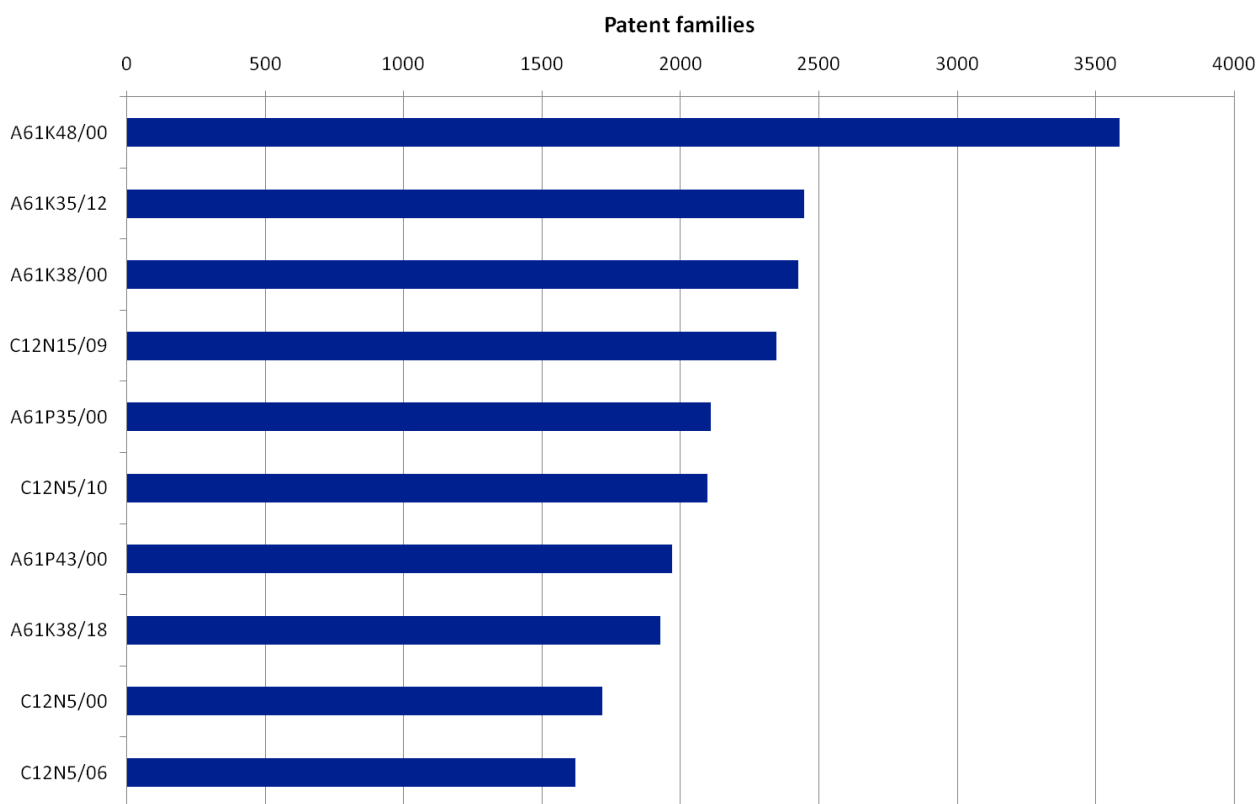


Figure 11: Top IPC sub-groups

Table 2: Key to IPC sub-groups referred to in Figure 11

A61K48/00	Medicinal preparations containing genetic material which is inserted into cells of the living body to treat genetic diseases; Gene therapy
A61K35/12	Medicinal preparations containing material or reaction products thereof with undetermined constitution -> Materials from mammals or birds
A61K38/00	Medicinal preparations containing peptides
C12N15/09	Mutation or genetic engineering; DNA or RNA concerning genetic engineering, vectors, e.g. plasmids, or their isolation, preparation or purification; Use of hosts therefor -> Recombinant DNA-technology
A61P35/00	Antineoplastic agents
C12N5/10	Undifferentiated human, animal or plant cells, e.g. cell lines; Tissues; Cultivation or maintenance thereof; Culture media therefor -> Cells modified by introduction of foreign genetic material, e.g. virus-transformed cells
A61P43/00	Drugs for specific purposes, not provided for in groups A61P0001000000-A61P0041000000
A61K38/18	Medicinal preparations containing peptides -> Peptides having more than 20 amino acids; Gastrins; Somatostatins; Melanotropins; Derivatives thereof -> from animals; from humans -> Growth factors; Growth regulators
C12N5/00	Undifferentiated human, animal or plant cells, e.g. cell lines; Tissues; Cultivation or maintenance thereof; Culture media therefor
C12N5/06	Undifferentiated human, animal or plant cells, e.g. cell lines; Tissues; Cultivation or maintenance thereof; Culture media therefor -> Animal cells or tissues

### 3 The UK landscape

#### 3.1 Top UK applicants

Figure 12 shows the top UK-based applicants within the regenerative medicine dataset. Note that the UK follows the international trend in having a large academic input with a number of universities and spin-out companies present in this list. Again, there is potential for further growth and development of regenerative medicine in the UK with no clear incumbent companies or large portfolios.

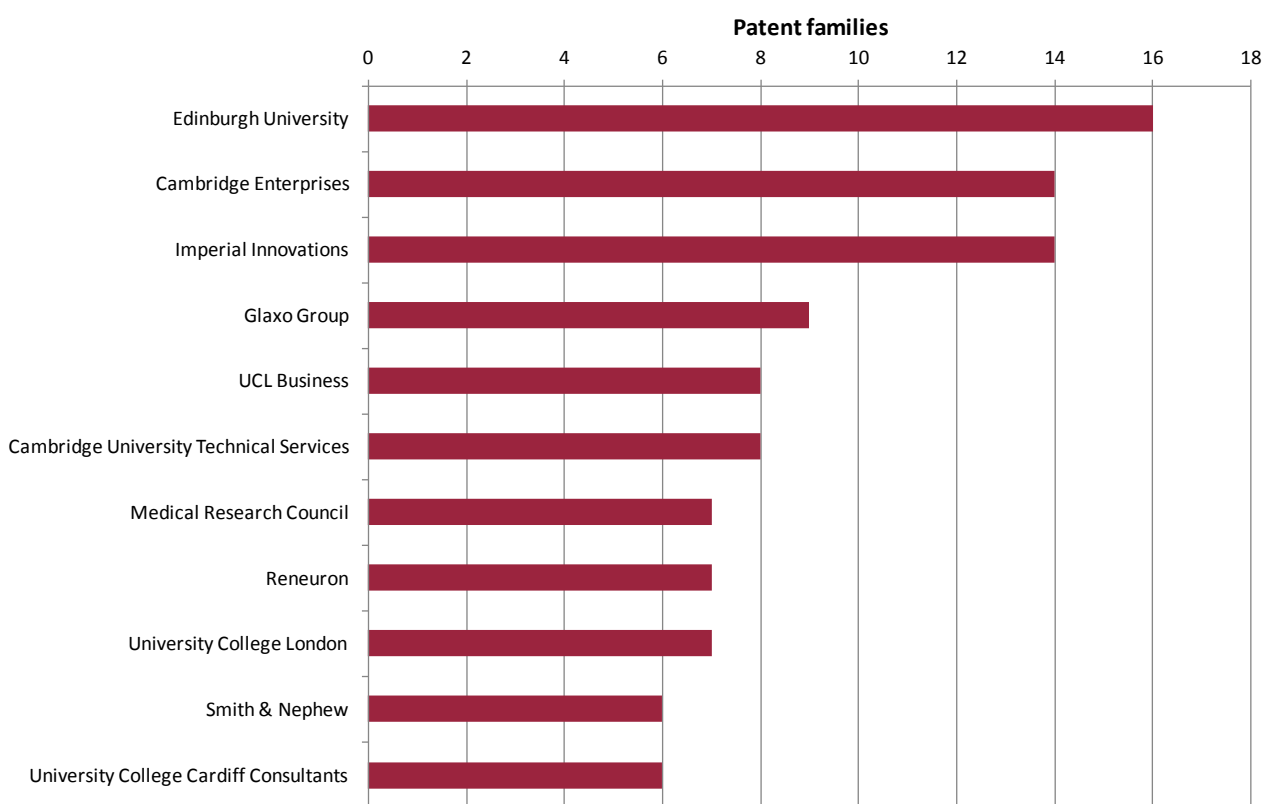


Figure 12: Top UK applicants

### 3.2 UK inventor mobility

Figure 13 shows the top worldwide applicants with named UK inventors on their published patents. This list is largely a reflection of Figure 12 since all but two entries in the list are UK-based organisations. The exceptions are Ares Trading, based in Switzerland, and Smithkline Beecham Corporation, which appears to be a historical entity dating from before the merger forming the current Glaxo Group. Nevertheless, this entity used a USA-based address in its patents but with inventors based in the UK.

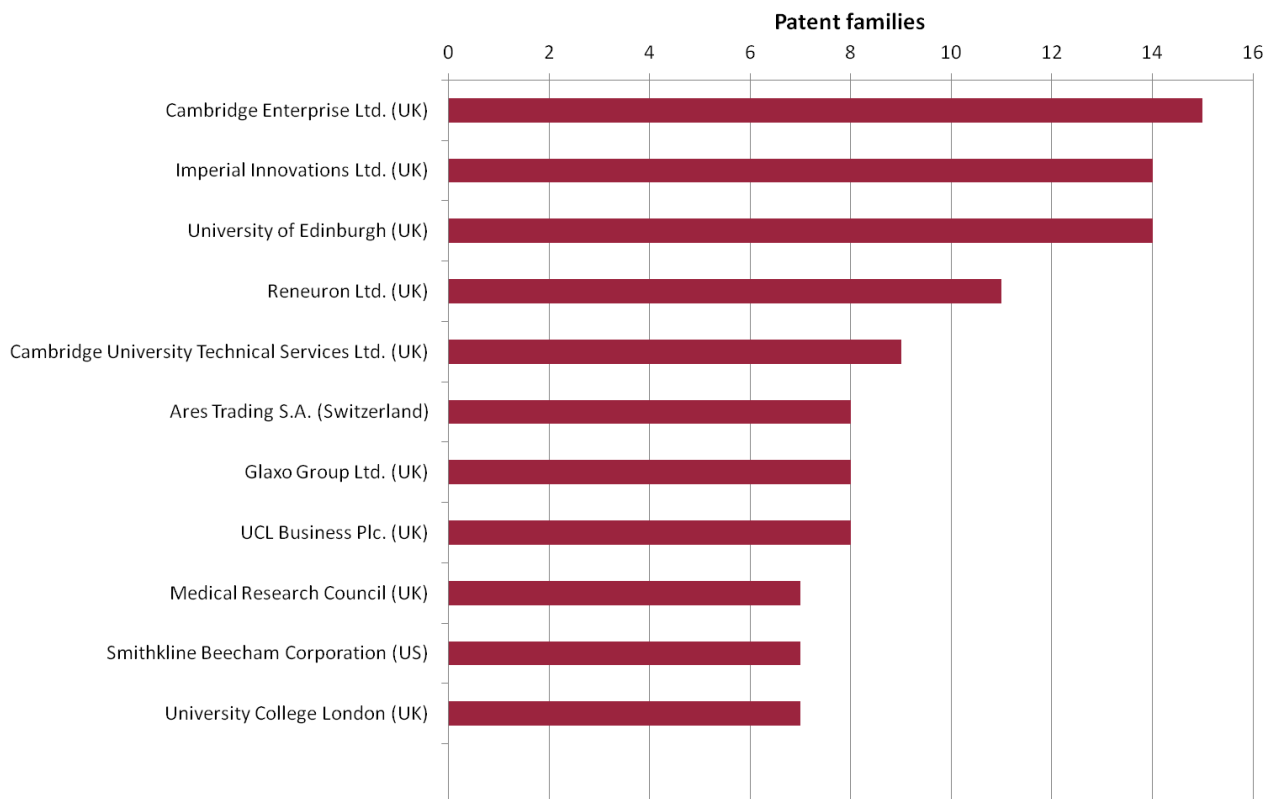


Figure 13: Top worldwide applicants with named UK-based inventors

### 3.3 How active is the UK?

A subset of the main worldwide dataset designed to reflect all UK patenting activity was selected. Figure 14 shows the year-on-year change in UK patenting activity against the worldwide year-on-year change in regenerative medicine patenting shown in Figure 2; this shows that changes in UK patenting activity in regenerative medicine are generally lower or more negative than changes worldwide, and there is a clearer declining trend than worldwide patenting activity.

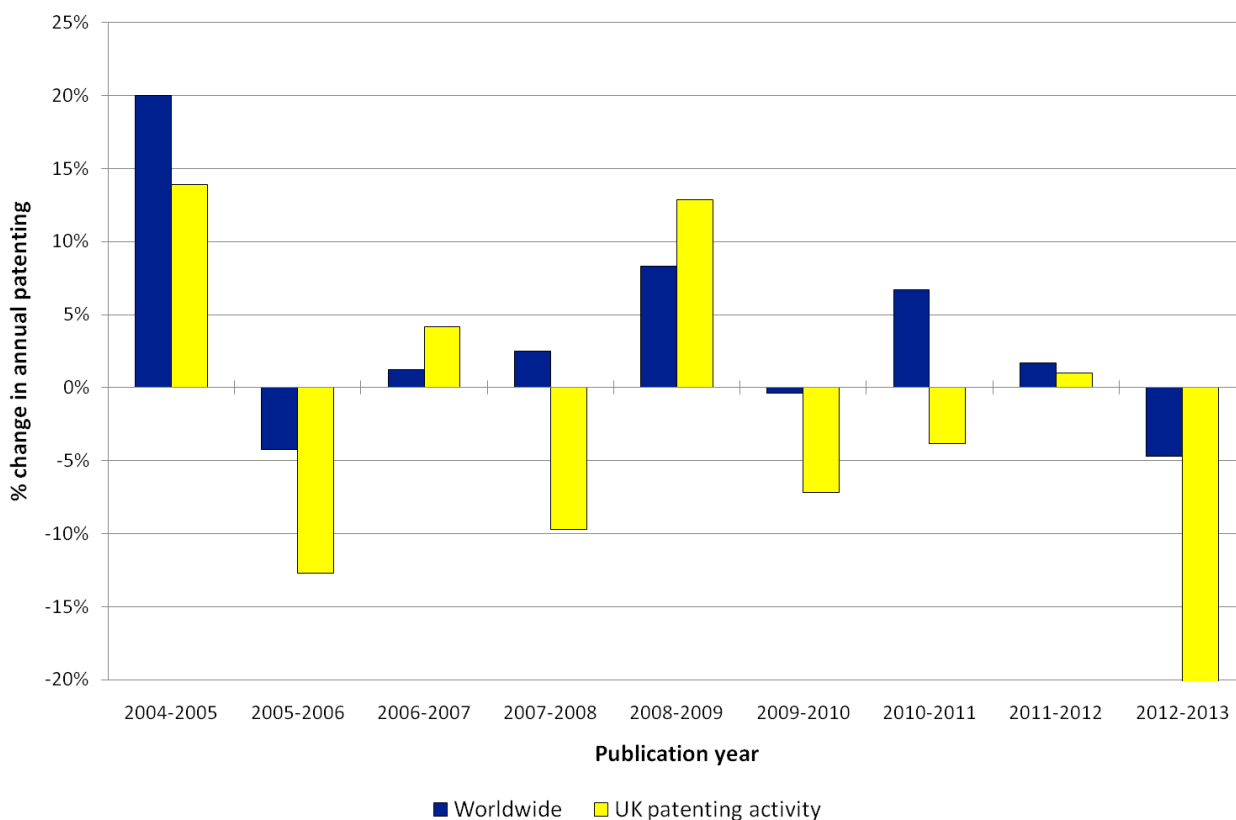


Figure 14: Year-on-year change in UK and worldwide regenerative medicine patenting

Similar patent subsets were created to reflect patenting activity taking place in several comparator countries (France, Germany, USA, Japan, China, and Korea) to produce the comparison chart shown in Figure 15.

It is notable that Korea, China, and to some extent Japan, exhibit consistent growth each year over the period. Although these countries all showed a low specialisation, there is still growth which outstrips other comparator countries according to this measure, and so the emergence of regenerative medicine appears to be focussed in these places. The UK has seen a greater decline between 2012-2013 than the comparator countries.

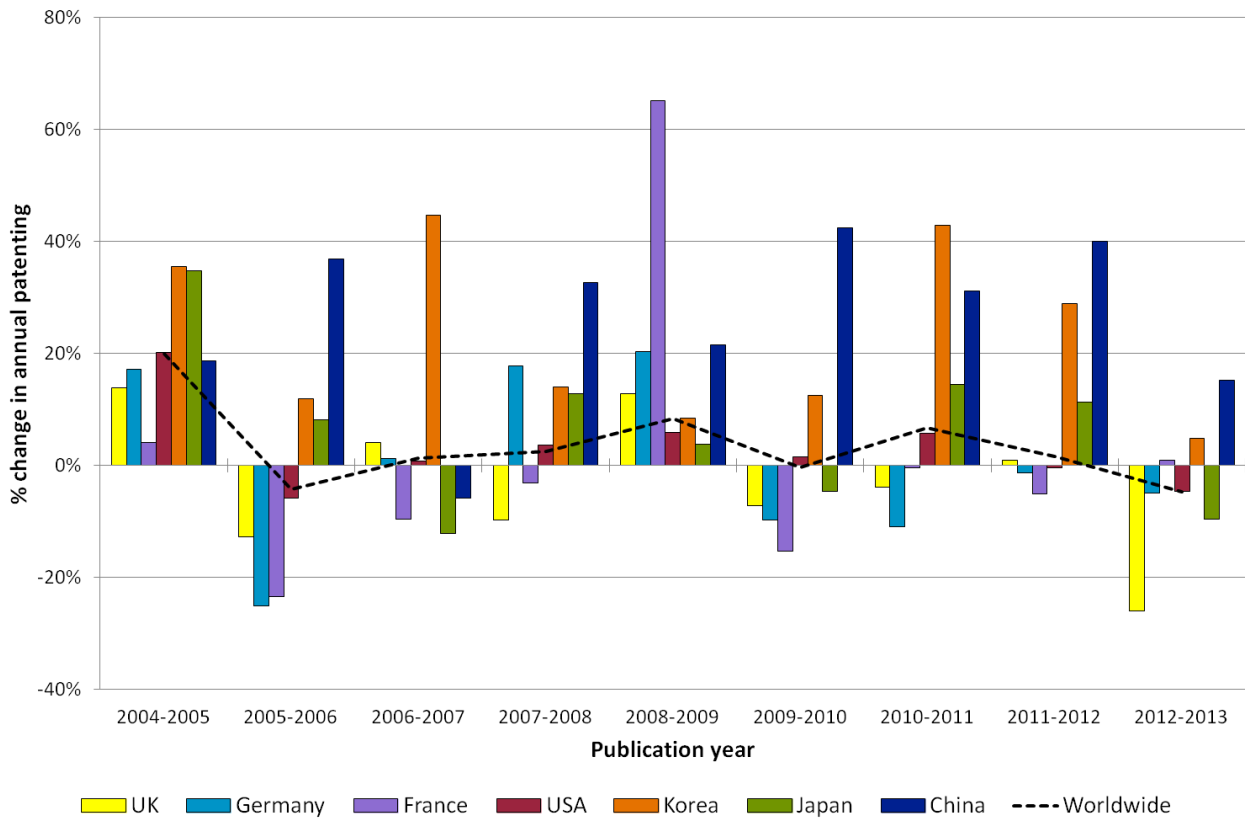


Figure 15: Year-on-year change in UK regenerative medicine patenting activity against comparison countries

## 4 Patent landscape map analysis

In order to give a snapshot as to what the patent landscape looks like for this technology space, a patent map provides a visual representation of the dataset. Published patents (not patent families) are represented on a patent map by dots and the more intense the concentration of patents (*i.e.* the more closely related they are) the higher the topography as shown by contour lines. The patents are grouped according to the occurrence of keywords in the title and abstract and examples of the reoccurring keywords appear on the patent map<sup>5</sup>.

The landscape map for regenerative medicine is shown in Figure 16. The major topics in the map are *antibodies*, *nucleic/nucleic acid*, *amino acids*, *vector*, *polypeptide*, *DNA*, *Seq ID*, and *host*. The major peaks in this map are annotated *polypeptide*, *amino acid*, *polynucleotide* (top left), *vector*, *virus*, *nucleic acid* (bottom right), *CNS*, *sclerosis*, *nestin* (bottom), and *prostate*, *antibody*, *colon* (centre right).

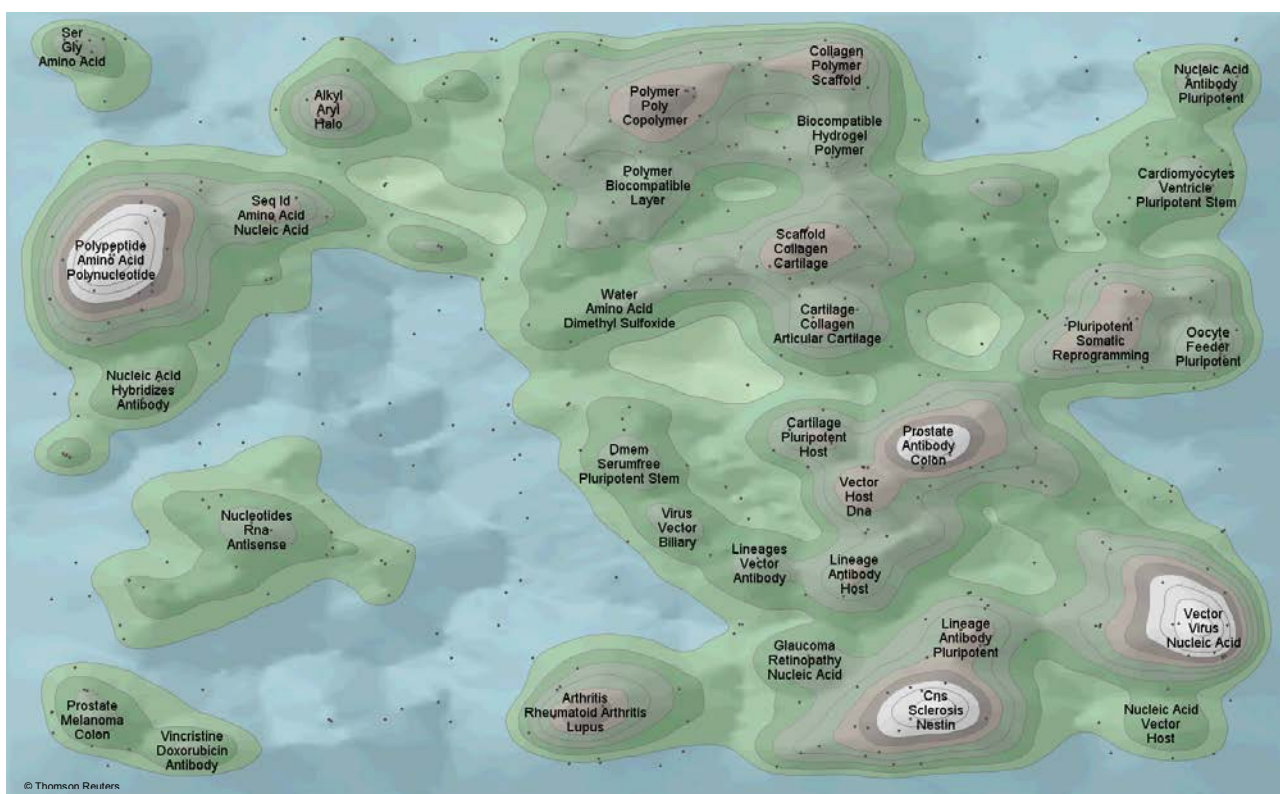


Figure 16: Patent landscape map of patent claims, abstract, and title terms

<sup>5</sup> Further details regarding how patent landscape maps are produced is given in Appendix C.



Upon investigation, it is found that the leading applicants are distributed fairly homogeneously throughout the map. This indicates the applicants have a wide and varied interest over the entire field and have not yet settled into distinct niches, as might be expected once the science base is developed by businesses into competing products and services; this is consistent with the high level of collaboration which occurs in regenerative medicine.

Figure 17 shows an alternative landscape map formed using words only from the “use” and “advantage” parts of the DWPI abstract. In this case the following topics relating to particular conditions are found: cancer, multiple sclerosis, autoimmune, (rheumatoid) arthritis, anemia, prostate cancer, lupus, lymphoma, carcinoma, colon cancer, asthma, melanoma, ovarian cancer, sarcoma, multiple myeloma, retinopathy, allergy, dementia, cervical cancer, and epilepsy.

In this map the applicants are, again, distributed homogeneously, with the exception of Kyoto University which has a cluster around the peak annotated *reprogrammed*, *nanog*, *nucleic acid* at the right, and these patents are highlighted in red. This region of the map appears to relate to methods of reprogramming and the differentiation of pluripotent stem cells into various cell types.

Other peaks in the map include the bottom right, *sclerosis*, *lateral*, *oligodendrocyte*. There is a concentration of patents here relating to the differentiation of neural and neuronal stem cells for the treatment of neurodegenerative conditions. Conditions in this region are epilepsy, *multiple sclerosis*, *cerebral palsy*, and *dementia*. A further large peak, or twin peak, is towards the left, annotated *amino acid*, *seq ID*, *nucleotide*, and *specifically binds*, *seq*, *nucleic acid*. This region is focussed on new polypeptides and polynucleotides for screening, diagnosis, prevention, and treatment of various conditions.

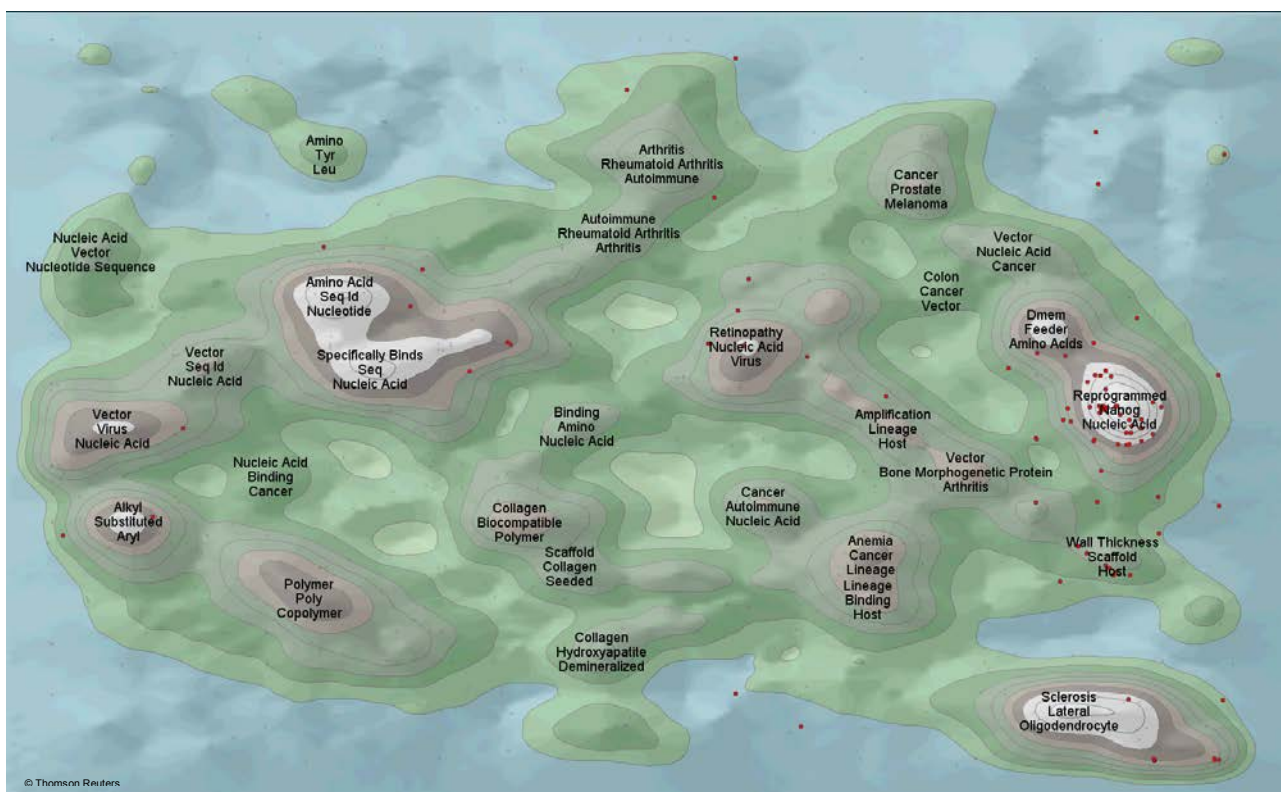


Figure 17: Patent landscape map of advantages and uses



## 5 Conclusions

Regenerative medicine patenting has seen a slowdown and perhaps even a decline over a ten year period from 2004 to 2013, both worldwide and even more so in the UK. This slowdown contrasts with the general increase in patenting globally over the same period. Despite this, however, regenerative medicine also shows the patenting characteristics of an emerging technology with the potential for much further growth.

Almost two-thirds (64%) of regenerative medicine patent families have a first filing in the USA, with the next largest country being Japan at 9%. The UK is at 3%. Patent applicants based in the USA, however, account for just 50% of patent families, with other applicants spread across a diversity of countries, indicating the importance of patenting regenerative medicine in the USA for applicants worldwide.

An index of relative specialisation of applicant countries indicates that China, Korea, Japan, and Germany in particular show a low specialisation in regenerative medicine. However, China, Korea, and Japan show consistent and positive percentage growth since 2004 and so the emergence of regenerative medicine may turn out to be focussed in those countries. Israel, Australia, Canada and the USA, meanwhile, have the greatest levels of specialisation. The UK has a small degree of specialisation in this study although this has increased since an earlier study on regenerative medicine patenting which was carried out in 2011.

Sixteen of the top twenty applicants in regenerative medicine are based in the USA, eight of them are academic institutions, and a further four are government departments or national research agencies. The University of California (USA) takes the leading position with 171 patent families, with General Hospital Corporation (USA) being some way behind at 102 patent families, and the University of John Hopkins (USA) in third place with 97 patent families. 28% of patent families overall include an academic applicant. The role of business in regenerative medicine patenting is therefore small and could be expected to increase as the scientific research base is subsequently developed and commercialised. The mix of top applicants includes newer entries and some who have become less active in recent years, suggesting that there is still a significant turnover of organisations involved with regenerative medicine and much more scope for organisations to establish and maintain leading positions. In the UK, Edinburgh University (16 patent families), Cambridge Enterprises Ltd. (14), and Imperial Innovations Ltd. (14) take the lead. UK inventors tend to be associated with UK applicant companies and universities.

The significant academic input into regenerative medicine and the lack of corporate incumbents has allowed a collaborative environment to flourish both in the US and across the globe, with significant levels of co-assigned patents between organisations. Collaborations from university to university, and university to business, have created a tightly integrated web of collaborations amongst the leading applicants. In the UK, separate clusters have developed but the levels of collaboration are still significant. Cambridge University, University College London, the University of Edinburgh, and associated technology transfer companies, have links to UK businesses and to the USA.

Patent landscape analysis further demonstrates the collaborative nature of regenerative medicine patenting, with applicants tending to show interest in the entire field and to lack niches. Inventions addressing various types of cancer and neurodegenerative conditions are found to be common topics within the landscape.

# Appendix A Interpretation notes

## A.1 Patent databases used

The *Thomson Reuters World Patent Index (WPI)* was interrogated using *Thomson Innovation*<sup>6</sup>, a web-based patent analytics tool produced by *Thomson Reuters*. This database holds bibliographic and abstract data of published patents and patent applications derived from the majority of leading industrialised countries and patent organisations, e.g. the World Intellectual Property Organisation (WIPO), European Patent Office (EPO) and the African Regional Industry Property Organisation (ARIPO). It should be noted that patents are generally classified and published 18 months after the priority date. This should be borne in mind when considering recent patent trends (within the last 18 months).

The WPI database contains one record for each patent family. A patent family is defined as all documents directly or indirectly linked via a priority document. This provides an indication of the number of inventions an applicant may hold, as opposed to how many individual patent applications they might have filed in different countries for the same invention.

## A.2 Priority date, application date and publication date

**Priority date:** The earliest date of an associated patent application containing information about the invention.

**Publication date:** The date when the patent application is published (normally 18 months after the priority date or the application date, whichever is earlier).

Analysis by priority year gives the earliest indication of invention.

## A.3 WO and EP patent applications

International patent applications (WO) and European patent applications (EP) may be made through the World Intellectual Property Organization (WIPO) and the European Patent Office (EPO) respectively.

International patent applications may designate any signatory states or regions to the Patent Cooperation Treaty (PCT) and will have the same effect as national or regional patent applications in each designated state or region, leading to a granted patent in each state or region.

European patent applications are regional patent applications which may designate any signatory state to the European Patent Convention (EPC), and lead to granted patents having the same effect as a bundle of national patents for the designated states.

Figures for patent families with WO and EP as priority country have been included for completeness although no single attributable country is immediately apparent.

---

<sup>6</sup> <http://info.thomsoninnovation.com>

## A.4 Patent documents analysed

The regenerative medicine patent dataset for analysis was identified in conjunction with patent examiner technology-specific expertise. A search strategy was developed and the resulting dataset was extracted in April 2014 using International Patent Classification (IPC) codes, Co-operative Patent Classification (CPC) codes and keyword searching of titles and abstracts in the *Thomson Reuters World Patent Index (WPI)* and limited to patent families with publications from 2004 to 2013.

The applicant and inventor data was cleaned to remove duplicate entries arising from spelling errors, initialisation, international variation (Ltd, Pty, GmbH *etc.*), or equivalence (Ltd., Limited, *etc.*).

## A.5 Analytics software used

The main computer software used for this report is a text mining and analytics package called *VantagePoint*<sup>7</sup> produced by *Search Technology* in the USA. The patent records exported from *Thomson Innovation* were imported into *VantagePoint* where the data is cleaned and analysed. The patent landscape maps used in this report were produced using *Thomson Innovation*.

---

<sup>7</sup> <http://www.thevantagepoint.com>

## Appendix B Relative Specialisation Index

Relative Specialisation Index (RSI) was calculated as a correction to absolute numbers of patent families in order to account for the fact that some countries file more patent applications than others in all fields of technology. In particular, US and Japanese inventors are prolific patentees. RSI compares the fraction of regenerative medicine patents found in each country to the fraction of patents found in that country overall. A logarithm is applied to scale the fractions more suitably. The formula is given below:

$$\log_{10} \left( \frac{n_i/n_{total}}{N_i/N_{total}} \right)$$

where

$n_i$  = number of regenerative medicine patents in country  $i$

$n_{total}$  = total number of regenerative medicine patents in dataset

$N_i$  = total number of patents in country  $i$

$N_{total}$  = total number of patents in dataset

The effect of this is to highlight countries (in this study, Israel and Australia in particular, as shown in Figure 5) which have a greater level of patenting in regenerative medicine than expected from their overall level of patenting, and which would otherwise languish much further down in the lists, unnoticed.

## Appendix C Patent landscape maps

A patent landscape map is a visual representation of a dataset and is generated by applying a complex algorithm with four stages:

- i)* **Harvesting documents** – When the software harvests the documents it reads the text from each document (ranging from titles through to the full text). Non-relevant words, known as stopwords, (e.g. “a”, “an”, “able”, “about” *etc*) are then discounted and words with common stems are then associated together (e.g. “measure”, “measures”, “measuring”, “measurement” *etc*).
- ii)* **Analysing documents** – Words are then analysed to see how many times they appear in each document in comparison with the words’ frequency in the overall dataset. During analysis, very frequently and very infrequently used words (*i.e.* words above and below a threshold) are eliminated from consideration. A topic list of statistically significant words is then created.
- iii)* **Clustering documents** – A Naive Bayes classifier is used to assign document vectors and Vector Space Modelling is applied to plot documents in n-dimensional space (*i.e.* documents with similar topics are clustered around a central coordinate). The application of different vectors (*i.e.* topics) enables the relative positions of documents in n-dimensional space to be varied.
- iv)* **Creating the patent map** – The final n-dimensional model is then rendered into a two-dimensional map using a self-organising mapping algorithm. Contours are created to simulate a depth dimension. The final map can sometimes be misleading because it is important to interpret the map as if it were formed on a three-dimensional sphere.

Thus, in summary, patents are represented on the patent map by dots and the more intense the concentration of patents (*i.e.* the more closely related they are) the higher the topography as shown by contour lines. The patents are grouped according to the occurrence of keywords in the title and abstract and examples of the reoccurring keywords appear on the patent map. Please remember there is no relationship between the patent landscape maps and any geographical map.

Please note that the patent maps shown in this report are snapshots of the patent landscape, and that patent maps are best used an interactive tool where analysis of specific areas, patents, applicants, inventors *etc* can be undertaken ‘on-the-fly’.

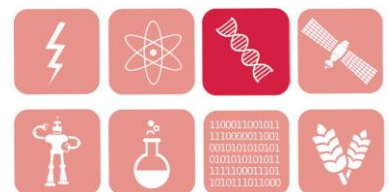






Concept House  
Cardiff Road  
Newport  
NP10 8QQ  
United Kingdom

[www.ipo.gov.uk](http://www.ipo.gov.uk)



#8Great