



政策研究大学院大学
NATIONAL GRADUATE INSTITUTE
FOR POLICY STUDIES

政策研究大学院大学 科学技術イノベーション政策研究センター ワーキングペーパー (SciREX-WP)
National Graduate Institute for Policy Studies, Science for RE-Designing Science,
Technology and Innovation Policy Center (SciREX Center) Working Paper

[SciREX-WP-2022-#01]

**論文・特許クラスター分析を用いたファイトケミカルの学術研究・
技術開発動向調査**

Investigation of Academic Research and Technology Development Trends
of Phytochemicals Using Paper and Patent Cluster Analysis

2022/07

Tatsuo Sasaki (National Graduate Institute for Policy Studies)

佐々木達郎(政策研究大学院大学 専門職)



SciREX Center
WORKING PAPER

政策研究大学院大学

科学技術イノベーション政策研究センター (SciREX センター)
ワーキングペーパー SciREX-WP-2022-#01

[SciREX-WP-2022-#01]

論文・特許クラスター分析を用いたファイトケミカルの学術研究・技術開発動向調査

Investigation of Academic Research and Technology Development Trends of Phytochemicals Using Paper and Patent Cluster Analysis

2022年7月

佐々木 達郎

政策研究大学院大学SciREXセンター 専門職

※. 本ワーキングペーパーの著作権は、著者もしくは政策研究大学院大学科学技術イノベーション政策研究センターに帰属しています。本ワーキングペーパーに含まれる情報を、個人利用の範囲を超えて転載、またはコピーを行う場合には、政策研究大学院大学科学技術イノベーション政策研究センターによる事前の承諾が必要となりますので、以下までご連絡ください。

【連絡先】政策研究大学院大学科学技術イノベーション政策研究センター
TEL: 03-6439-6318 / E-Mail: scirex-center@grips.ac.jp

目次

Abstract	4
エグゼクティブサマリー	5
1 イントロダクション	6
2 分析手法.....	7
3 分析結果.....	7
3.1 分析データの抽出.....	7
3.2 母集団の二次元可視化.....	9
3.3 基本統計.....	16
3.4 時系列変化.....	21
3.5 国・地域別、機関別集計.....	28
4 まとめ	43
謝辞	44
参考文献	44

Abstract

We extracted 44,712 papers and 7,486 patents related to phytochemicals and visualized them by clustering analysis using text data such as titles and abstracts. The results show that the papers and patents are divided into seven areas: "Plant Medicine and Treatment," "Plant Hormone and Gene Expression," "Biosynthesis Control," "Molecular Structure Analysis and Identification," "Green Manufacturing," "Genetic Resources," "Plant Culture," and "Herbal Medicine and Herbal Medicine."

The number of papers and patent applications has been increasing at a constant rate, and research on "plant hormones," "endophytic bacteria," and "growth-promoting bacteria" has begun since 2011. In 2016 and beyond, "phytohormones," "plant-microbe interactions," "growth-promoting bacteria," and "green synthesis (metal nanoparticles)," which are similar to basic technologies, have seen a marked increase. Green Synthesis (metal nanoparticles) are increasing significantly.

In 2020, we believe that we are still in the stage of deepening basic understanding, but we have identified areas where technological development is progressing in linkage with academic research, such as "growth-promoting bacteria" and "plant-microbe interactions."

エグゼクティブサマリー

ファイトケミカルに関連する論文44,712件と特許7,486件を抽出し、タイトル・アブストラクト等のテキストデータを用いたクラスタリング分析による可視化を行った。【植物による医薬・治療】【植物ホルモン・遺伝子発現】【生合成制御】【分子構造分析・同定】【グリーン製造】【遺伝資源】【植物培養】【薬草・漢方薬】の7領域に分かれている様子が確認された。

論文数・特許出願数は一定の割合で増加しており、2011年以降から「植物ホルモン」「内生菌」「成長促進細菌」などの研究が始まっている。また応用先としての「癌治療」「抗炎症作用」「糖尿病治療」や、補助的な計測技術である「NMR構造解析」「HPLC」が増加している。2016年以降では基礎技術に近い「植物ホルモン」「植物-微生物相互作用」「成長促進細菌」「グリーン合成(金属ナノ粒子)」の増加が著しい。

2020年段階において基礎的な理解を深める段階と思われるが、「成長促進細菌」や「植物-微生物相互作用」等の学術研究とリンクして技術開発が進んでいる分野が確認された。

論文・特許クラスター分析を用いたファイトケミカルの学術研究・技術開発動向調査

1 イン트로ダクション

科学技術・イノベーション(STI: Science, Technology and Innovation)が経済・社会システムの中でどのような影響をもたらしているかを明らかにすることは、政策と研究の両面から重要な課題である。特に、第6期科学技術・イノベーション基本計画では、人文・社会科学も含む「総合知」の概念も導入され、科学技術・イノベーションと経済社会システムの関係性を理解し、政策に実装することが重要な課題となっている。

我々は2011年に開始された科学技術・イノベーション(STI)政策における「政策のための科学」推進事業 (Science for RE-designing Science Technology and Innovation Policy: SciREX事業)の中で科学技術・イノベーション政策の経済社会効果評価を中核的な課題と位置付けて研究に取り組んできた。2021年度からは共進化実現プログラムの研究プロジェクト(科学技術・イノベーション政策の経済社会効果分析の政策形成プロセスへの実装)にて、研究開発投資戦略の文書の内容(科学技術シナリオ)を経済モデルに投入するパラメータに変換するためのプロトコル開発に取り組んでいる。

経済パラメータへの変換を行うためには、着目する研究開発分野において、①どのような学術研究がなされているか、②学術研究からどのような技術が生まれているのか、③新しい技術を採用した新製品・サービスとして社会に普及すると産業にどのような影響を及ぼすのか、の3つの要素を明らかにすることが重要である。我々は研究開発投資に関わる政策文書や論文・特許データを用いてこれらの関係性を求め、技術ロードマップの形で整理した上で経済モデルベースのシミュレーター(黒田ら、2016; CRDS, 2016)に投入していく予定である。

本稿では上記①②の要素を明らかにするため、近年注目が集まっているファイトケミカル技術を事例として採用し、論文・特許データを用いて分析した結果を示す。具体的にはファイトケミカルの要素技術ごとに学術研究および技術開発動向を可視化し、国・地域・機関ごとの特徴を明らかにするため、論文・特許の俯瞰分析を実施した結果を報告する。

ファイトケミカルとは「一次代謝物と二次代謝物の両方を合わせ、植物が生産する全ての化合物」を指す(CRDS, 2021)。植物から生み出される化合物の生合成メカニズムを解明・制御することで、有効な成分・素材を高効率に生産することを目的としている。

このためには、生体分子・化合物を同定する化学系研究と、進化やゲノムを扱う生物学研究が連携することによってファイトケミカルの生合成メカニズムを解明し、プロセスに人為的な介入を行うことによって高効率な生産を実現する必要がある。さらに、異なる植物や微生物の酵素を組み合わせることで新たな生合成経路をデザインすることで、天然植物とは異なる新たな化合物の生産も可能となる。このように生まれた新たな植物由来成分は、健康食品や医薬品の原料として市場への展開が期待されている(CRDS, 2021)。

分析の結果、ファイトケミカルに関しては要素技術が7領域に分かれており、論文・特許は一定の割合で増加していた。特許については天然由来植物成分を使った健康機能性食品や漢方薬の技

術が多かったが、近年は学術研究とリンクする技術も増加している様子がうかがえた。学術研究の成果は癌治療等の医療へ応用につながるものが増加しており、今後も論文数が増加して研究が活発になっていくことが見込まれた。

2章で分析手法の流れを示し、3章で分析データについて解説を行い、4章で全体の総括を行った。

2 分析手法

科学技術・イノベーション投資が経済効果を生み出すまでのプロセスを検討する上で、“新製品・サービスの普及が産業に及ぼす影響は、採用されている技術の質によっても大きく変動する”という性質は重要な意味を持つ。すなわち研究開発分野の全体像に加えて、要素技術毎に学術研究・技術開発の動向を明らかにしていく必要がある。

そこで本調査においてはファイトケミカルに関連する論文・特許を分析対象として抽出し、自然言語処理によって類似する文献群(クラスター)を作成して集計・可視化を行った。

主な分析の流れを下記に示す。

1. 分析データの抽出: 検索キーワードを設定し、論文(Web of Science: 1980年-2020年収録2021年6月調達)および特許(TotalPatentOne)を用いて検索することで関連する論文・特許を抽出した。
2. 対象データの2次元可視化: 論文・特許のアブストラクト等のテキストデータを用いてトピックモデル(LDA: Latent Dirichlet Allocation; 潜在的ディリクレ配分法)による類似度評価を行い、トピック配分が類似する文献が近くに配置されるように2次元可視化することで研究・技術俯瞰図を作成した。
3. 基本統計: 抽出した論文・特許全体の文献数推移および国別文献数推移を作成した。
4. 時系列変化: 全体の年次推移を参考にして期間を分割し、期間毎に研究・技術俯瞰図を作成した。文献数および文献数の伸び率に基づいてヒートマップを作成し、研究開発が盛んな領域の可視化を行った。
5. 国・地域別、機関別統計: 論文・特許の国・地域別、機関別文献数推移を集計し、それぞれの研究・技術俯瞰図を作成した。

3 分析結果

3.1 分析データの抽出

戦略プロポーザル「ファイトケミカル生成原理とその活用のための研究開発戦略 ～未利用植物資源から革新的価値を創出する学術基盤の創成～」(CRDS, 2021)の記述および専門家からのコメントを参考として検索キーワードを設定し、論文(Web of Science)、特許(TotalPatentOne)のデータベースにて検索を実施した。検索条件と検索された文献数を表1に示す。ファイトケミカルに関連する論文のうち、重複・酷似・記載の短いものを除外して58,629件、特許をファミリーの重複・英語機械翻訳欠損・記載の短いものを除外して18,415件抽出した。

表 1 ファイトケミカル 母集団検索式と件数

No.	分類	WoS正規表現*	WoS件数	特許検索式†	特許件数
1	phytochemical	phyto.{0,2}(chemical polymer)	23357	TAC:(phytochemical OR phytopolymer OR (phyto PRE1 (chemical* OR polymer)))	1173
2	medicinal plants	medicinal.{0,2}(plant herbs)	21052	TAC:(medicinal PRE1 (plants OR plant OR herbs OR herb))	16054
3	plant biosynthesis	plant[~\$]*?\$(~\$+?) {0,5}bio.{0,2}synthesis bio.{0,2}synthesis[~\$]*?\$(~\$+?) {0,5}plant	3961	TAC:(plant NEAR5 (biosynthesis* OR (bio PRE0 synthesis*)))	1551
4	plant hormones	plant.{0,2}hormone	4391	TAC:((plant PRE0 hormone*) OR plant hormone*)	2433
5	plant secondary metabolites	plant[~\$]*?\$(~\$+?) {0,3}metabolite	3182	TAC:(plant PRE3 metabolite*)	550
6	plant genetic variation (植物未利用遺伝子を発掘する研究)	plant[~\$]*?\$(~\$+?) {0,3}genetic.{0,2}(variation resources diversity) genetic.{0,2}(variation resources diversity)[~\$]*?\$(~\$+?) {0,3}plant	2316	TAC(plant NEAR3 ((genetic PRE0 variation*) OR (genetic PRE0 resource*) OR (genetic PRE0 diversity*) OR (genetic PRE0 variation*)))	141
7	plant made pharmaceuticals (植物でワクチン等の医薬品を作る分野)	(plant phyto).{0,2}(based made derived) AND (pharmaceutical medicine drug)	1655	TAC:(((plant PRE0 based*) OR (plant PRE0 made*) OR (plant PRE0 derived*) OR (phyto PRE0 made*) OR (phyto PRE0 derived*)) AND (pharmaceutical* OR medicine* OR drug*))	3495
8	plant-microbe interaction (農業への応用、耐病性)	plant.{0,2}(microbe bacteria microorganisms)[~\$]*?\$(~\$+?) {0,2}interact interact[~\$]*?\$(~\$+?) {0,2}plant.{0,2}(microbe bacteria microorganisms)	1026	TAC:(((plant PRE0 microbe*) OR (plant PRE0 bacteria*) OR (plant PRE0 microorganism*)) NEAR2 interact*)	18
9	plant-insect interaction (農業への応用、昆虫食害耐性)	plant.{0,2}(insect pest)[~\$]*?\$(~\$+?) {0,2}interact interact[~\$]*?\$(~\$+?) {0,2}plant.{0,2}(insect pest)	605	TAC:(((plant PRE0 insect*) OR (plant PRE0 pest*)) NEAR2 interact*)	6

10	遺伝子導入、倍数体操作、遺伝子発現制御	plant AND (gene.{0,2}transfer ploidy.{0,2}control gene.{0,2}expression.{0,2}control)	2711	TAC:(plant* AND ((gene PRE0 transfer*) OR (ploidy PRE0 control*) OR (gene PRE0 expression PRE0 control*)))	1083
----	---------------------	--	------	--	------

*Web of Scienceで全ての分類に共通する論文検索条件は「2001 ≤ pubyear ≤ 2020、doctype1=Article or Proceedings Paper、has_abstract=Y」である。

†TotalPatent Oneで全ての分類に共通する特許検索条件は、PG:(Application) AND PD:[2001-01-01 TO *] AND PC:(CN OR EP OR JP OR KR OR US OR WO OR CA OR DE OR FR OR GB OR TW)である。

3.2 母集団の二次元可視化

ファイトケミカルに関する論文58,629件について、タイトルとアブストラクトの文章を対象として、トピックモデルを用い1000次元弱のトピックを抽出した。その上で各テキストのトピックデータについて、多様体学習により2次元に圧縮して可視化した。図1中、赤破線楕円は筆頭トピックが共通である文献をまとめた。KW(Keyword)は、集積の源となった筆頭トピックに含まれる上位5キーワードを示す。更に目視により、関連する上位概念的技術を紫色の楕円で囲った。各領域の名称は、以下のデータの順で、参考にして決定した。①トピックのキーワード、②RAKE(Rapid Automatic Keyword Extraction)によるキープレーズ、③WoSのkeyword、④WoSのkeyword_plus、⑤タイトル、要約のテキスト。

なお、本稿では筆頭トピックが共通する文献群を赤破線で囲った領域キーワードを「」で、上位概念である紫色で囲った領域キーワードを【】で示す。

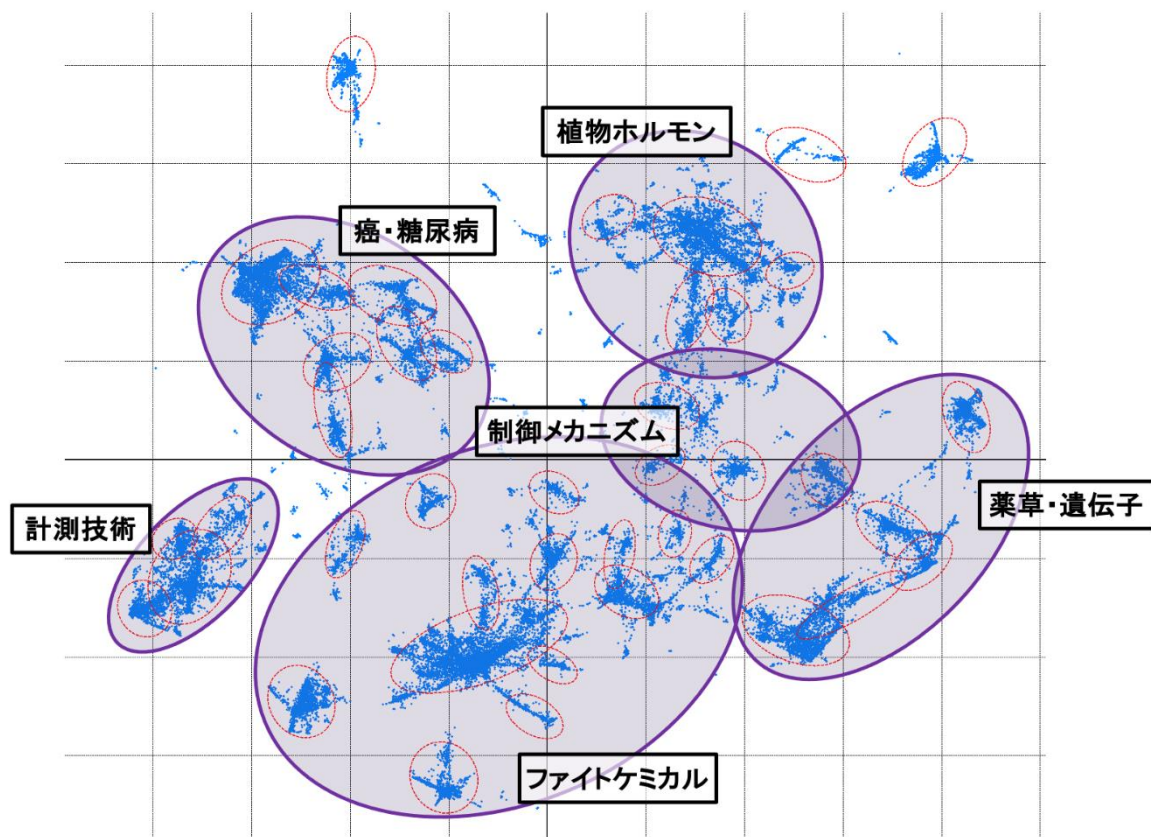
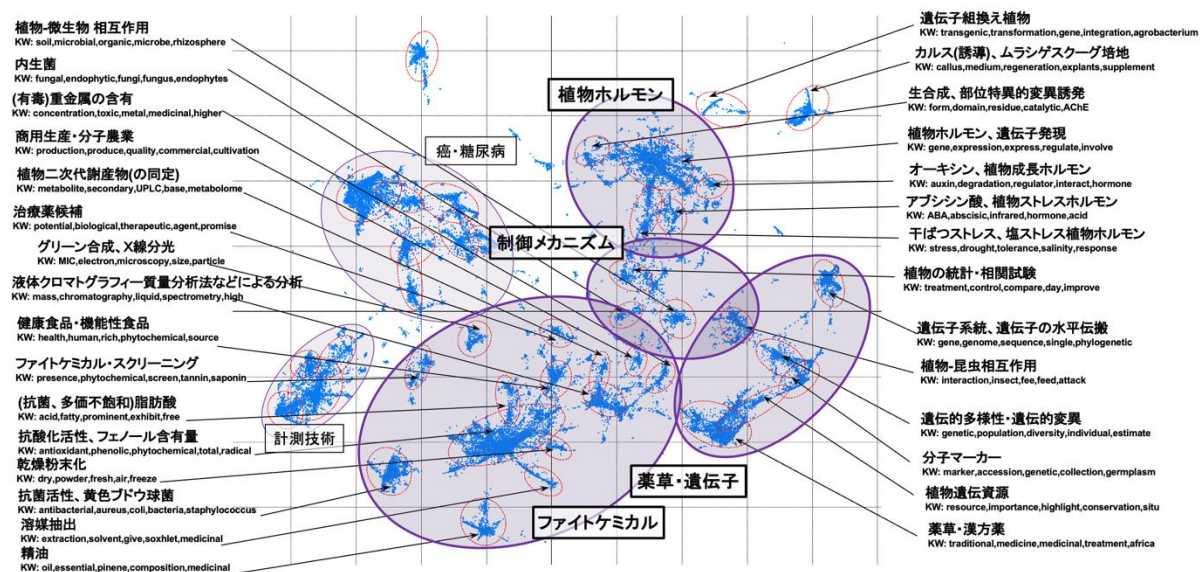


図 1 ファイトケミカル 論文トピックモデル計算の多様体学習による2次元可視化(ノイズ除去前)

論文に基づくファイトケミカルの領域は【ファイトケミカル】【薬草・遺伝子】【植物ホルモン】【制御メカニズム】【癌・糖尿病】【計測技術】の6つに分けられた。【癌・糖尿病】領域では、植物成分ではあるが癌や糖尿病に特化した研究が多く、病名で集積された可能性がある。



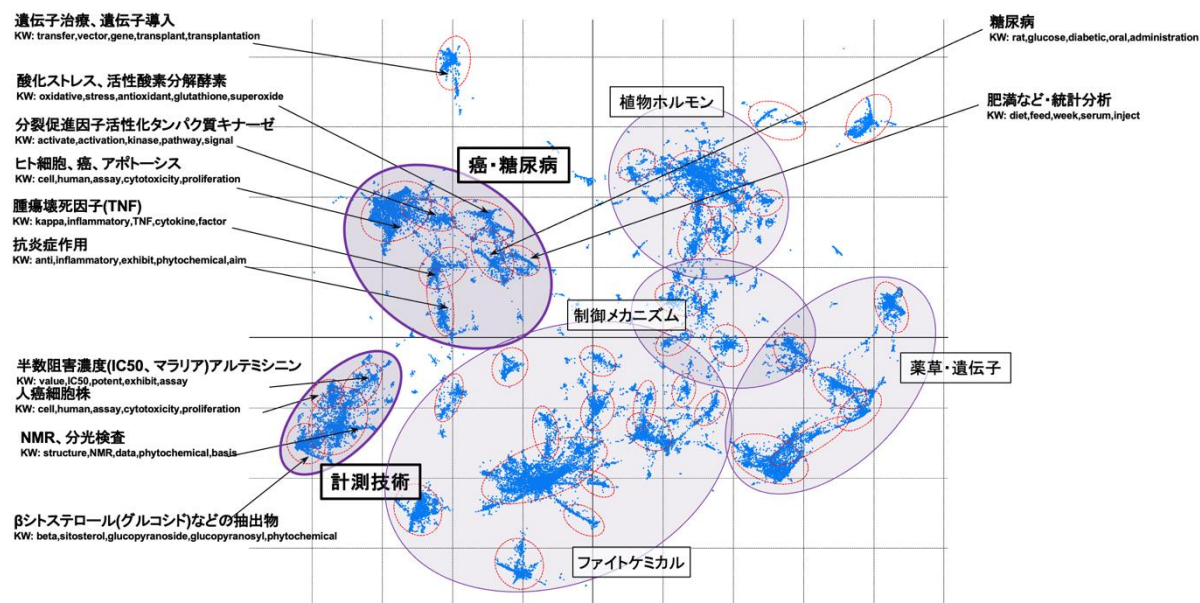


図2 ファイトケミカル 論文トピックモデル計算の多様体学習による2次元可視化 上段:ファイトケミカル】【薬草・遺伝子】【植物ホルモン】【制御メカニズム】領域詳細、下段:【癌・糖尿病】【計測技術】領域詳細

次に、キーワードに基づく検索式で抽出された論文内容を調査し、専門家を交えてファイトケミカルに関連する技術との適合性を確認した。「抗酸化活性、フェノール含有量」等のキーワードを含む領域(図3中ノイズ領域1)および「薬草・漢方薬」に関する領域(図3中ノイズ領域2)では植物性由来成分への言及があるものの、CRDS(2021)で言及しているファイトケミカル技術について記載された論文が少なく、今回の分析母集団から除外することとした。ただし、上記の領域に該当する論文のうちファイトケミカルに関する他のキーワード検索で該当した論文については分析母集団に加えた。ノイズと判定して母集団から削除することになった文献を図3に示す。緑円で囲まれた領域の論文を除外し、残った44,712件の論文を分析対象とした。

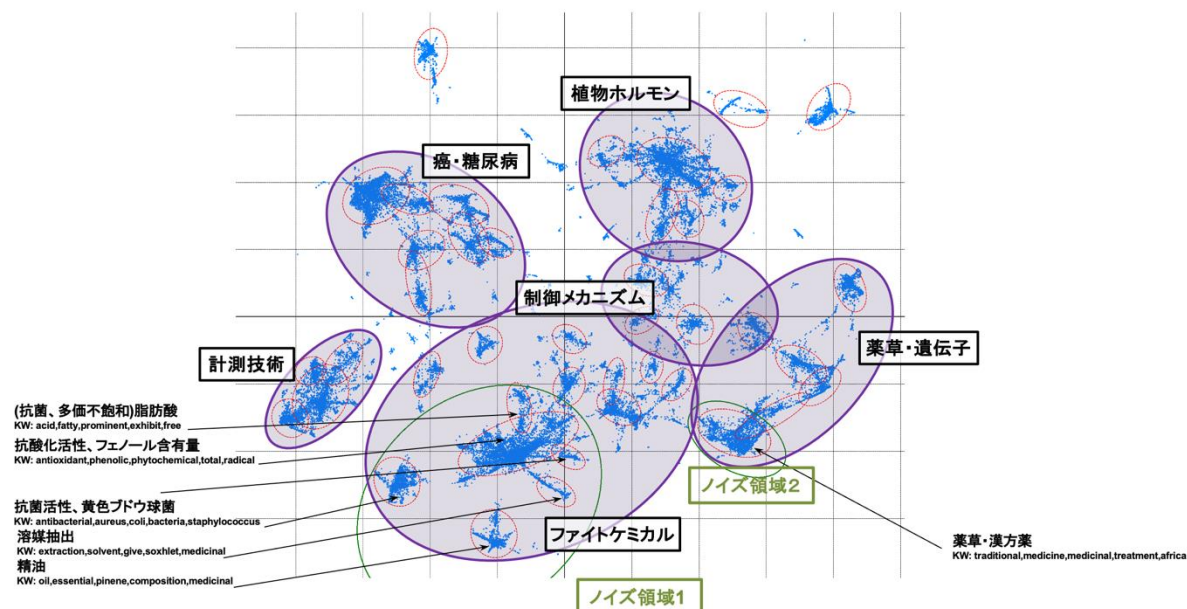


図3 ファイトケミカル 論文 トピックモデル計算の多様体学習による2次元可視化(ノイズ判定)

次にファイトケミカルに関する特許18,415件について、アブストラクトの文章を用いて、論文同様に分析を行った結果を図4に示す。

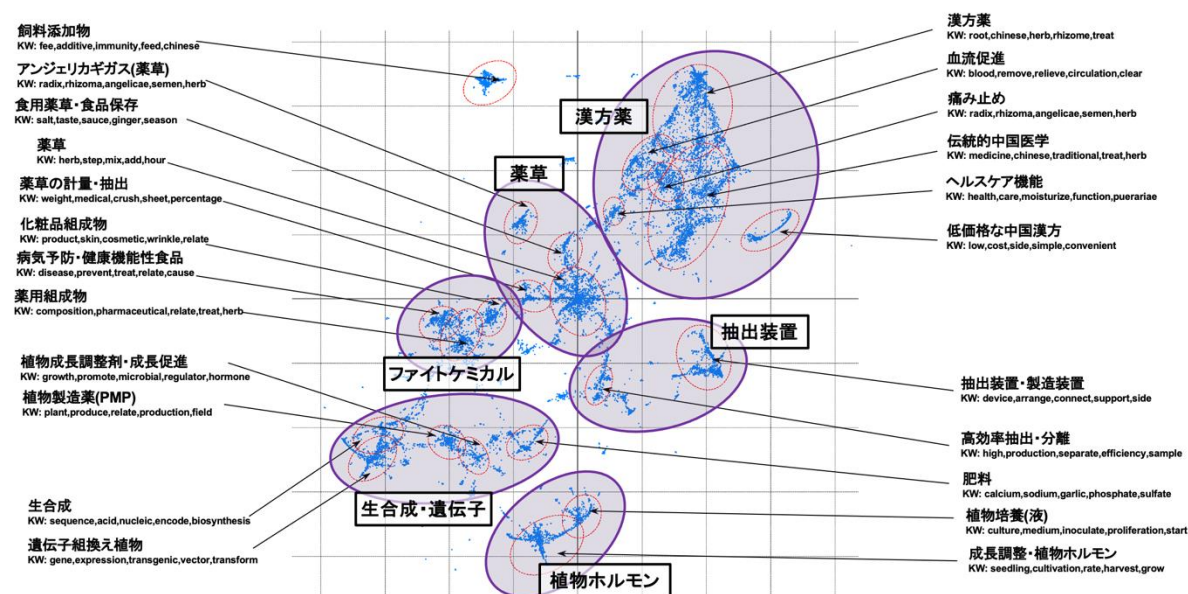


図4 ファイトケミカル 特許 トピックモデル計算の多様体学習による2次元可視化

【漢方薬】の領域において非常に特許数が多くなっている。論文にも「漢方薬」は含まれていたが、特許における割合が非常に大きい。論文と同様に特許に関してもキーワードに基づく検索式で抽出された内容を調査し、専門家を交えてファイトケミカルに関連する技術との適合性を確認した。【漢方薬】および【薬草】を含む領域(図5中ノイズ領域)では植物性由来成分への言及があるものの、CRDS(2021)で言及しているファイトケミカル技術について記載された特許が少なく、今回の分析母集団から除外することとした。ただし、除外した特許のうちファイトケミカル他キーワード検索

で該当する特許については分析母集団に加えることとした。ノイズ判定を行った結果を図4に示す。図中で赤と黄緑で示す特許を分析対象から除外し、更に年次上限を論文データと統一して2020年に限定を行った上で、最終的に残った7,486件の特許を分析対象とした。

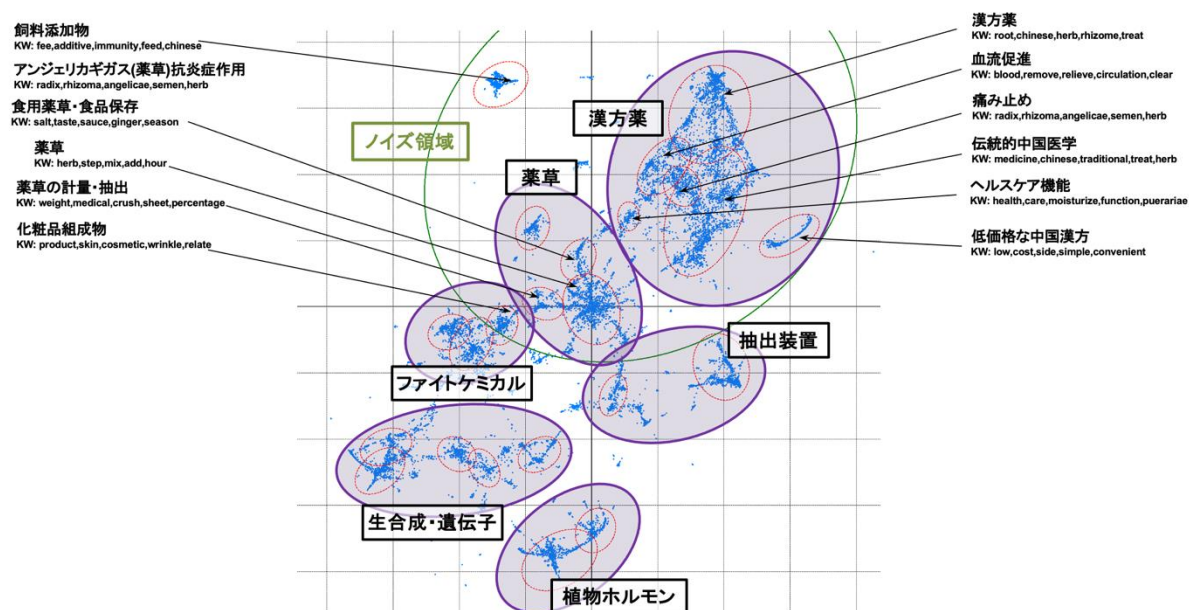


図 5 ファイトケミカル 特許トピックモデル計算の多様体学習による2次元可視化(ノイズ除去)

最後に、ファイトケミカルに関する論文44,712件と特許7,486件を合計した52,198件について、同様に分析した。結果を図6に、文献数のヒートマップを図7に示す。

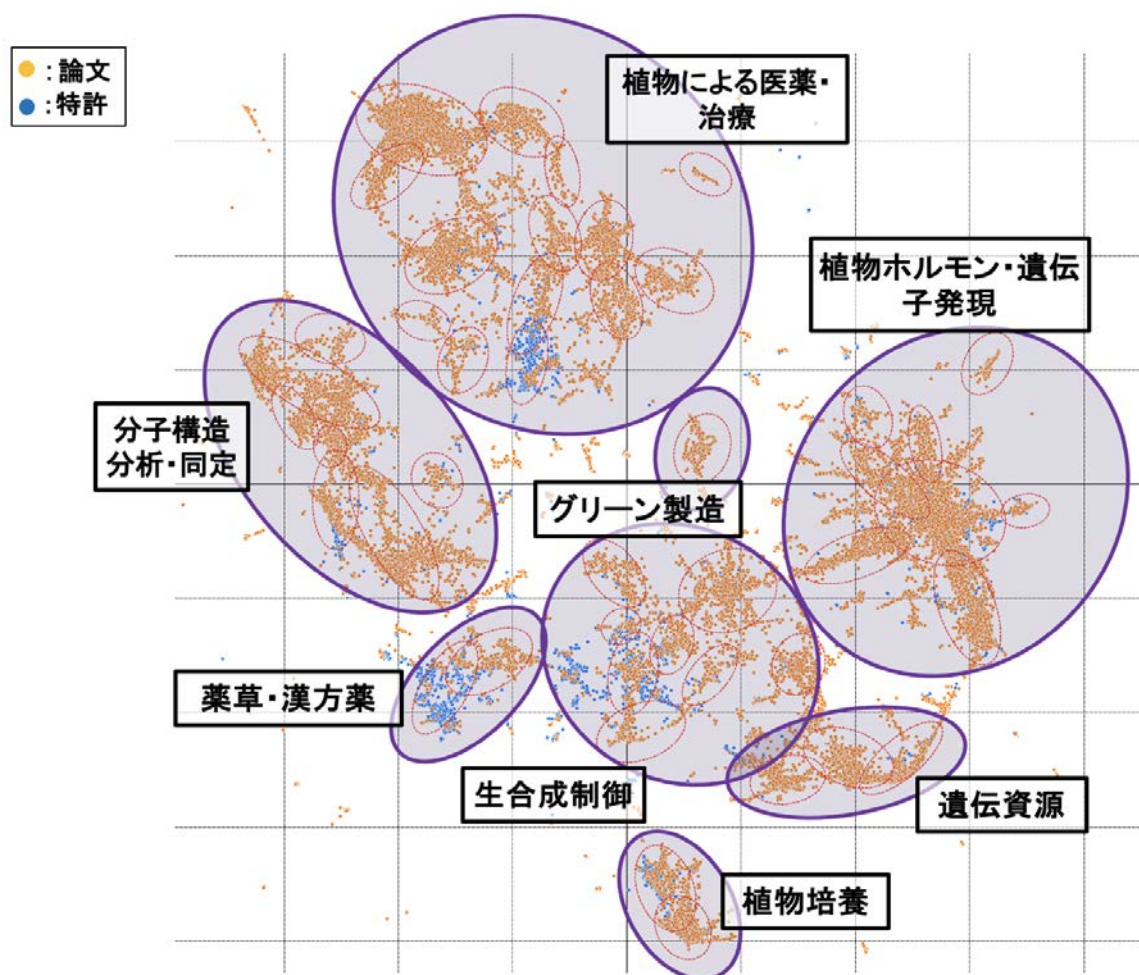


図 6 ファイトケミカル 論文+特許 トピックモデル計算の多様体学習による2次元可視化

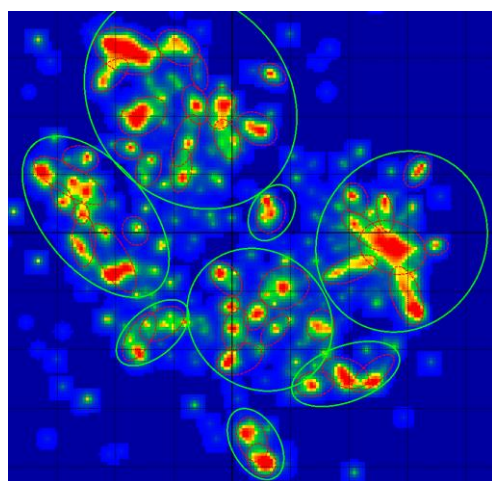


図 7 ファイトケミカル 論文+特許 トピックモデル計算の多様体学習による2次元可視化(ヒートマップ)



【植物による医薬・治療】の領域は主に各種病名に従って領域が分かれた。最大の領域は「癌治療」で、その他に「心臓疾患治療」「脳疾患治療」「糖尿病疾患治療」などに分かれた。なお、多くの特許が健康機能食品としての薬草として集積していた。【植物ホルモン・遺伝子発現】の領域は中心に「植物ホルモン、遺伝子発現」に関する領域が形成され、その周囲に「植物成長ホルモン」「ストレスホルモン」などの各種ホルモンや「遺伝子水平伝搬」などの領域が形成された。【生合成制御】領域には「内生菌」「植物-昆虫相互作用」「植物-微生物相互作用」などの生合成制御に関する技術が集積している。これら3つの領域の間に【グリーン製造】に関する「ナノ粒子生合成」に関する

領域が存在する。これらは俯瞰図の中心部に位置しており、ファイトケミカルの中心的な研究テーマと思われる。

【分子構造分析・同定】の領域には「NMRによる構造分析」「クロマトグラフィーによる質量分析」、更には各種組成の同定に関する技術が集積した。【遺伝資源】領域には「遺伝子多様性」「植物遺伝資源」「分子マーカー」に関する技術が集積している。【植物培養】の領域には「植物培養液」「カルス」に関する技術、【薬草・漢方薬】領域には「薬草」「中国漢方薬」に関する研究が集積している。

3.3 基本統計

ファイトケミカルに関する世界の論文・特許についての基本統計を以下に示す。なお、Web of Science(2021年6月調達)のデータにおいて、2020年分の論文収録状況に約1割程度の欠損が確認された。これは論文バルクデータ調達時期による収録の差異による影響と考えられる。特許はファミリー重複削除を行っているため、登録年はファミリー内の登録特許中の最古年を採用した。また、特許の国籍は発行国でなく、優先権主張先国とした。但しEPOに直接出願した特許などはEPのままとして特定の国に加算せずにEPとして集計した。

論文と特許を合計した文献数推移を図9に示す。ファイトケミカルに関連する文献総数は増加傾向を示している。特許(公開)も緩やかな増加を続けているが2018年を境に減少に転じている。これは中国での出願減少が原因であったが、他国では継続して出願されている。特許(公開)は2016年まで一定であったのがその後急激に増加している。特許(登録)数は特許(公開)数と比較すると件数が少なく、権利化はまだ活発ではない様子が伺える。

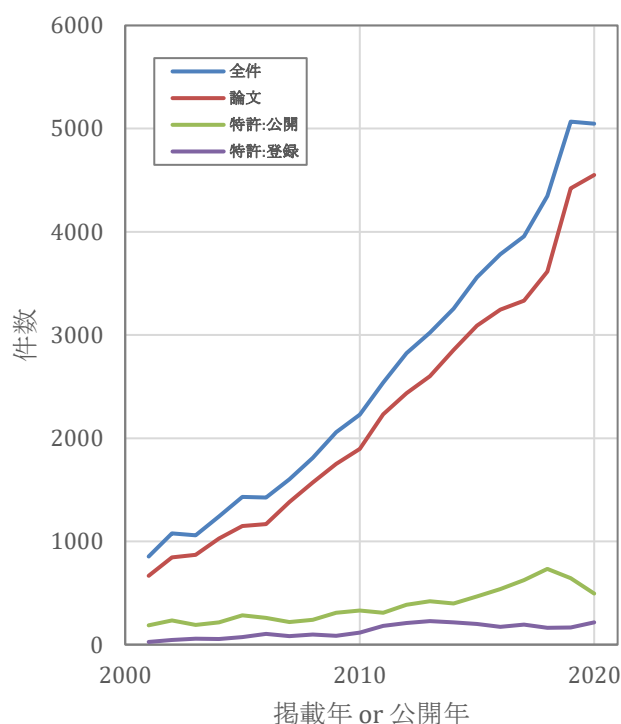


図9 ファイトケミカル 論文+特許 全体文献数推移

国・地域別の文献数を図10に示す。中国、米国、インドの三国の論文数が際立って多い。

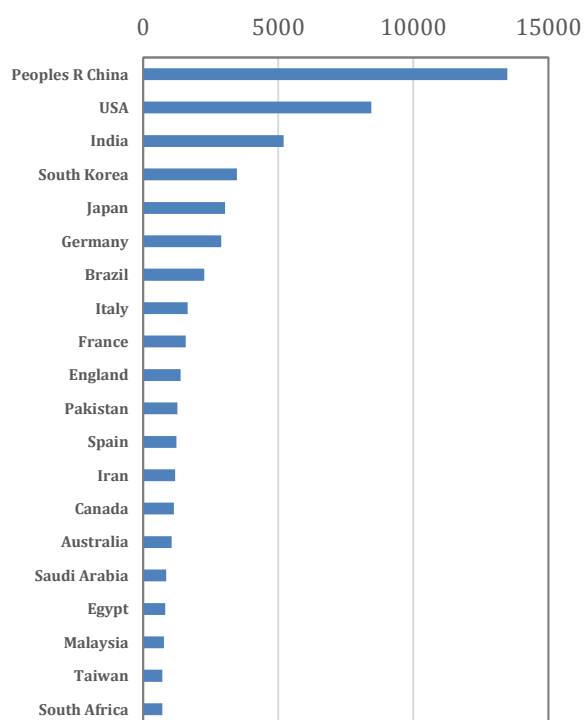


図 10 ファイトケミカル 論文+特許 国・地域別文献数

文献数上位6カ国の年次推移を図11に示す。欧米と日本が緩やかな増加を続けている一方で、中国、インド、韓国が高いペースで増加を続けており、特に中国の増加が著しい。

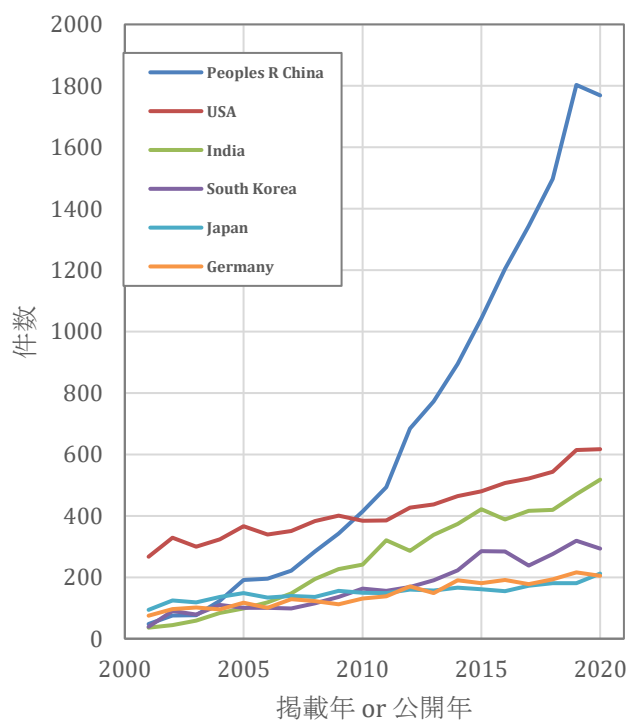


図 11 ファイトケミカル 論文+特許 国・地域別文献数推移(文献数上位6カ国)

論文・特許が集中している領域の差異を求めるため、ファイトケミカルについて、論文・特許の全体、論文のみ、特許のみの3パターンについて分布を求めた。ヒートマップによる可視化結果を図12に示す。

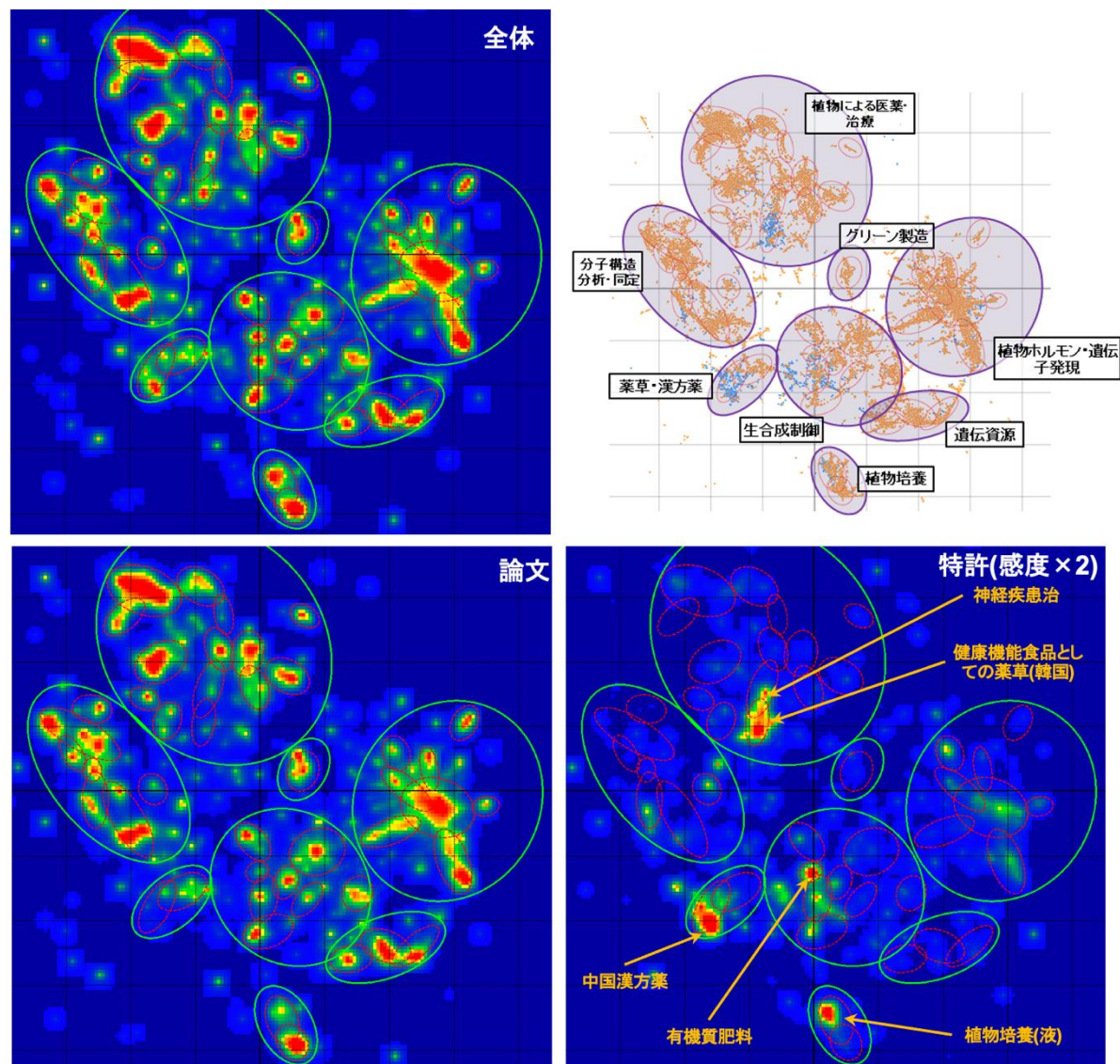


図 12 ファイトケミカル 論文+特許 トピックモデル計算の多様体学習による2次元可視化：論文、特許別分布

全体文件数の85.6%を論文が占めるため、全体分布と論文のみの分布はほぼ似た分布となった。特許は「中国漢方薬」「植物培養液」「健康機能食品としての薬草(韓国)」「有機質肥料」に多く集積している一方で、植物ホルモンなどの基礎研究に近い領域にも出願されていることが確認された。

一般的には俯瞰図の中心に位置する基礎的な理論および技術の領域は論文中心で特許が少なく、基礎技術を利用した応用技術が集まる周辺部の領域には特許が多く分布する階層構造が見られる。ファイトケミカルでは中心から少し離れた周辺部に特許が集まっている様子が確認された。

次に研究領域において学術研究(論文)と技術開発(特許)のどちらが活発に進んでいるかを可視化するため、俯瞰図を縦横0.5×0.5のメッシュに分割して、各メッシュ内の論文の割合を求めた。論文の割合を20%ごとに分割して可視化結果を図13に、論文割合を偏差値化して可視化した結果を図14に示す。

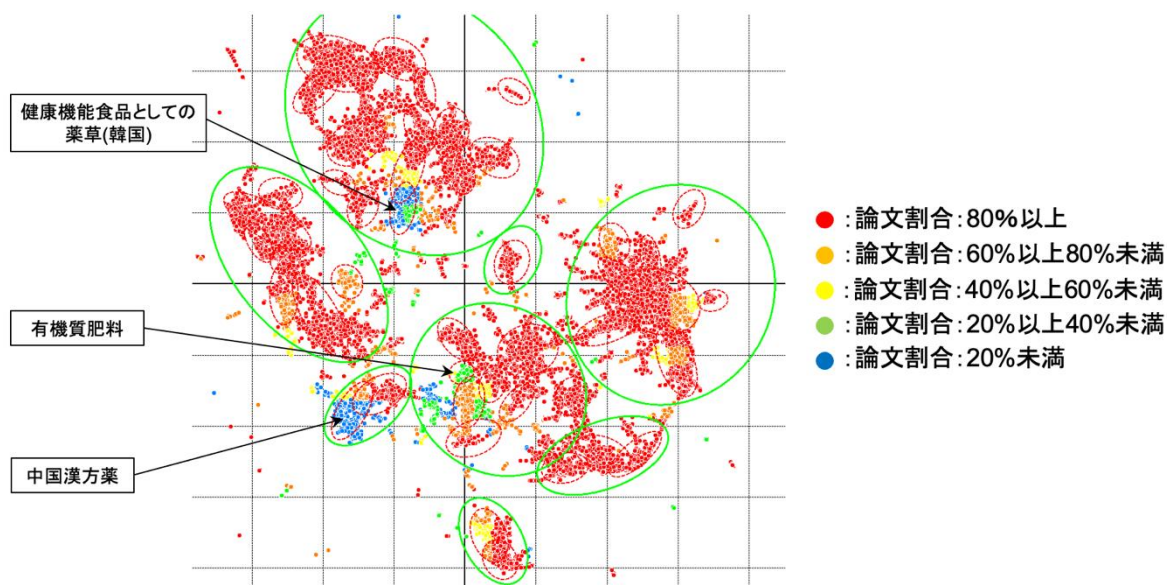


図 13 ファイトケミカル 論文+特許 : 論文と特許の割合分布(論文割合)

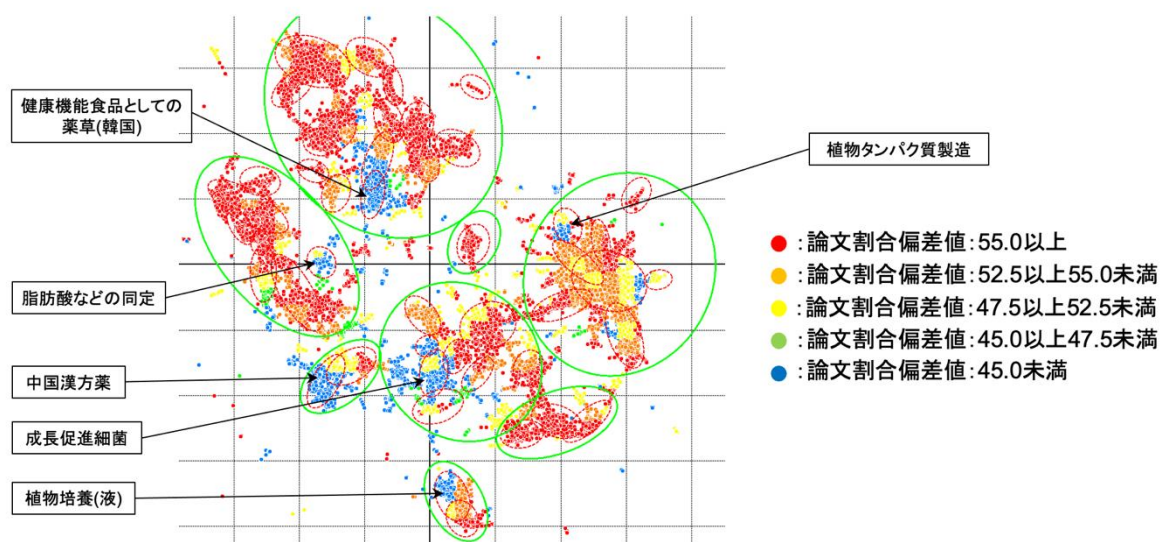


図 14 ファイトケミカル 論文+特許 : 論文と特許の割合分布(論文割合偏差値)

全体的にどの領域においても論文数の割合が高く、偏差値化した結果(図14)によって領域間の相対的な違いを可視化した。

俯瞰図の性質として中心に基本的概念が配置されるため学術研究が中心となって暖色系で示される。また、俯瞰図の最周辺は他領域とは異なる先端的な技術と考えられ、ここでも研究が中心で産業利用性が高くなく暖色で示される。その中間に基礎技術を応用する産業利用性が高い領域が

位置している。ファイトケミカルの基礎技術は俯瞰図中心にある【植物ホルモン・遺伝子発現】【生合成制御(の一部)】【遺伝資源】であり、その応用として【植物による医薬・治療】【薬草・漢方薬】が周辺に配置されている。

トピックモデルで計算された俯瞰図において、特定のトピックのウェイトが大きい論文が集まって密集領域を形成する。この局所的な集積の原因となる特徴的なトピックを主要トピックと呼ぶ。俯瞰図を縦横0.5×0.5のメッシュに分割し、各メッシュ内のトピック群に対して主要トピックが占める割合を算出し、各研究分野の独立性・関連性を評価した。主要トピック割合が高い領域は、他の領域では使われていない特徴的なトピックが集中的に存在していることになるため、独立性の高い技術と判断できる。一方で主要トピック割合が低い領域は、他の領域でも使われているトピックを多く含むことから、広範囲に応用されている適用先の広い技術と考えることができる。論文の主要トピック割合に応じて色分けしてプロットした結果を図15に示す。

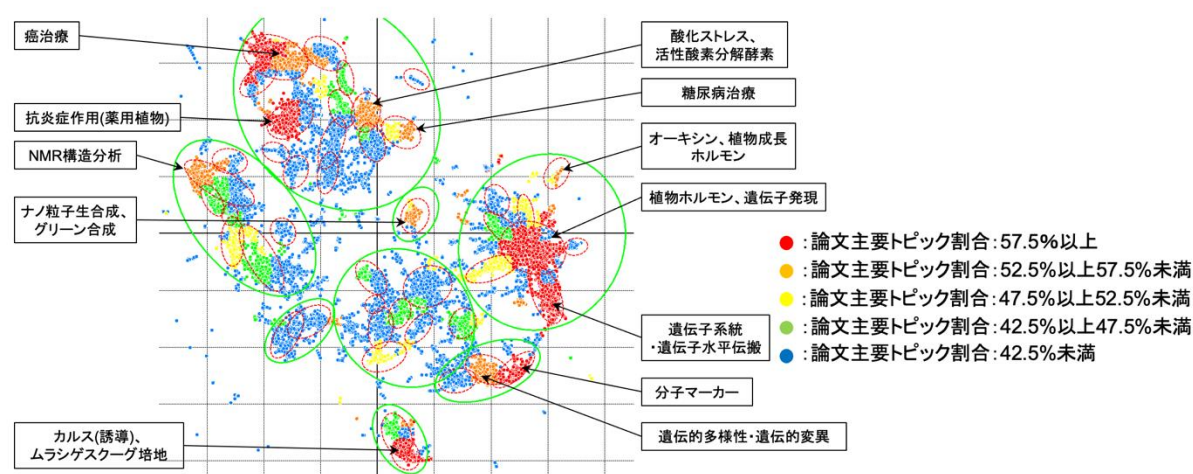


図 15 ファイトケミカル 論文+特許：論文主要トピック(ウェイト3.5%以上のトピック)割合プロット

一般的に、俯瞰図中心の領域には基礎技術が多く配置されており、周辺と共有するトピックが多いため寒色系となる。一方、中心から離れた領域では独自のトピックで集積するため主要トピック割合の高い暖色系となる傾向がある。ファイトケミカルもこの傾向に従って各技術が配置されていると考えられる。

次に特許に関して主要トピック割合に応じて色分けしてプロットした結果を図16に示す。各メッシュにおける主要トピック割合が低いほど、他領域においても広くトピックが使われており、応用への普及がなされていると解釈することができる。

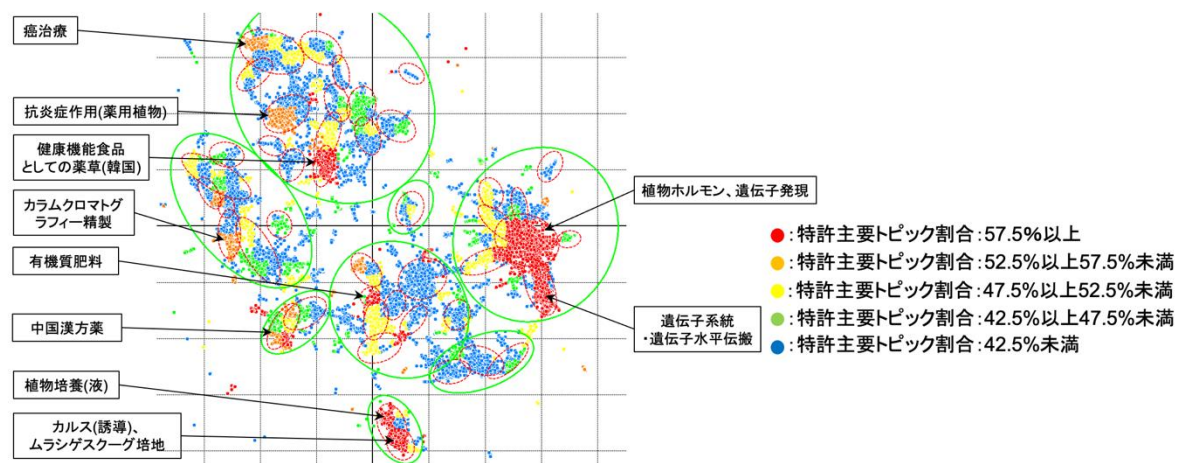


図 16 ファイトケミカル 論文+特許：特許主要トピック(ウェイト3.5%以上のトピック)割合プロット

全文献数に対して特許数の割合が低いため、分析精度は論文より劣る点に留意する必要があるが、論文で可視化した場合と同様の傾向が見られた。

3.4 時系列変化

ファイトケミカルの研究が時系列を追ってどのような分野で研究が進んできたかを明らかにするため、文献数の年次推移(図9)に基づいて5年ごとに分割して可視化した結果を図17に示す。

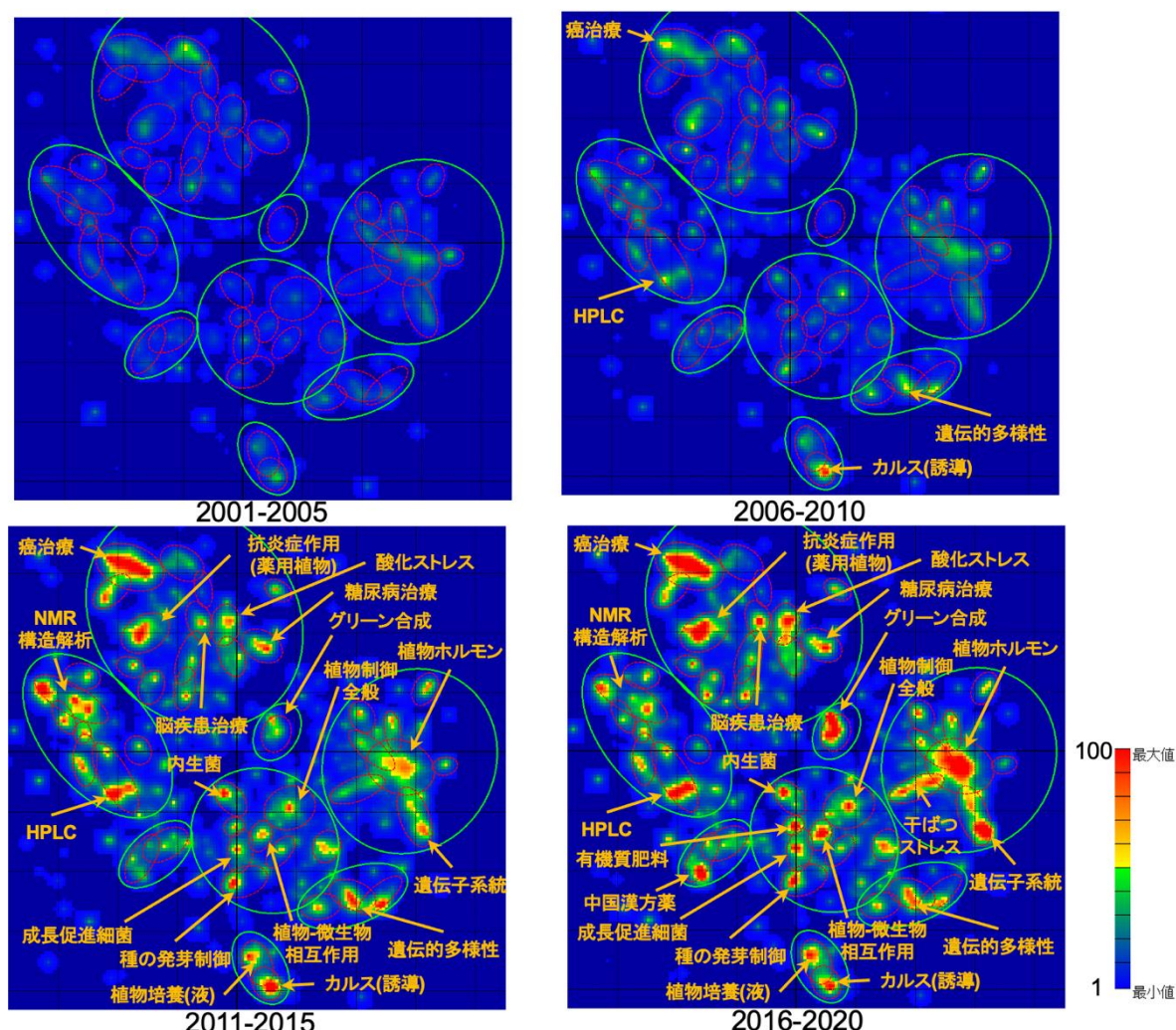


図 17 ファイトケミカル 論文+特許トピックモデル計算の多様体学習による2次元可視化：4期間ごと

2001から2010年までの10年間は俯瞰図周辺部の技術で件数がやや多く、ファイトケミカルの中心的な研究は少なかった。2011年以降から「植物ホルモン」「内生菌」「成長促進細菌」などの研究が始まっている。また応用先としての「癌治療」「抗炎症作用」「糖尿病治療」や、補助的な計測技術である「NMR構造解析」「HPLC」が増加している。2016年以降では補助的な計測技術がやや減少し、基礎技術に近い「植物ホルモン」「植物-微生物相互作用」「成長促進細菌」などの増加が著しい。また「グリーン合成(金属ナノ粒子)」も最近増加している。

次にファイトケミカルの論文に限定して図17と同様の可視化を行った結果を図18に、特許に限定して可視化を行った結果を図19に示す。

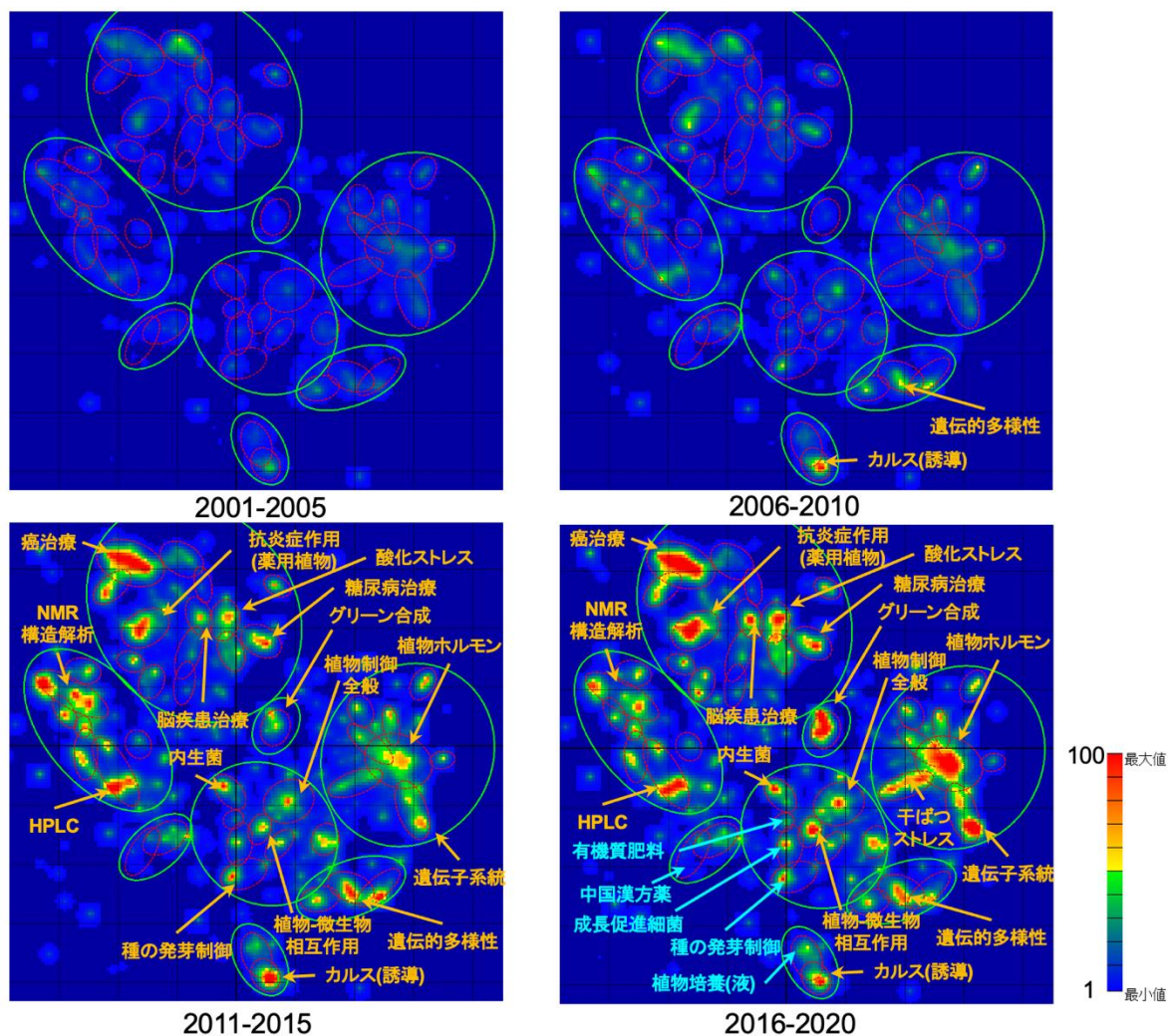


図 18 ファイトケミカル 論文 トピックモデル計算の多様体学習による2次元可視化 : 期間ごと

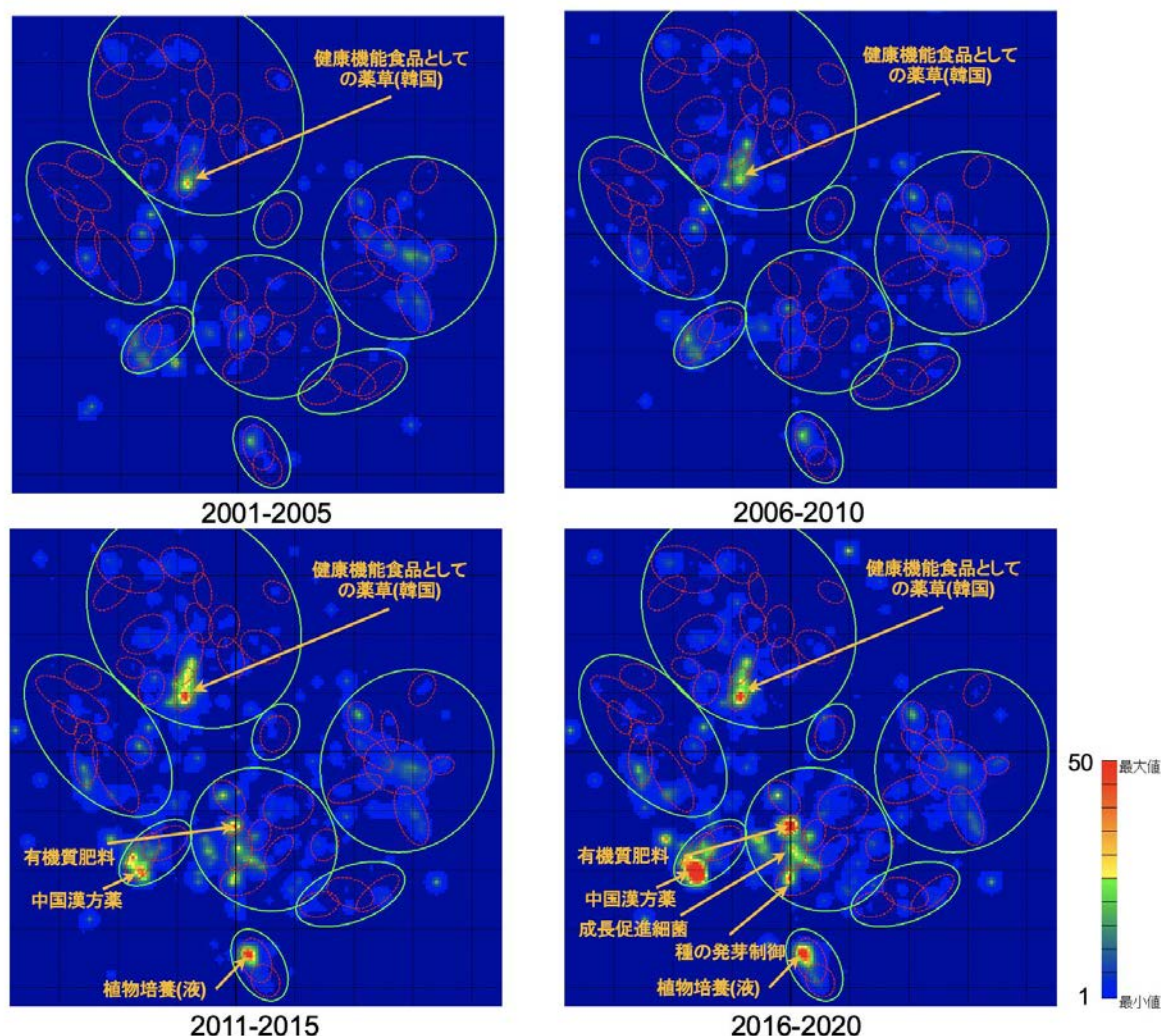


図 19 ファイトケミカル 特許 トピックモデル計算の多様体学習による2次元可視化：期間ごと

ファイトケミカルの特許は2016年以降に増加しているため、図18に示す論文のみで可視化した場合も2001年から2015年までの3期間の図は、全体の図(図17)と大きな違いがない。

図18中2016年以降のデータで、特許の寄与が無くなった領域を水色文字で示した。【生合成制御】内の「有機質肥料」「成長促進細菌」等で特許がなくなった分、ヒートマップが寒色よりになっている。一方、【植物ホルモン・遺伝子発現】は論文による研究主導が続いている。

特許のみでプロットした図19では、2001年当初から健康機能食品としての薬草が継続して出願されており、2011年からは中国漢方薬が急激に増え始めている。これらは論文の基礎研究より以前の時期から出願されているため、ファイトケミカルとの関連性が低い可能性がある。

研究開発が今後も活発に続き、論文や特許が増加しているエマージング領域を明らかにするため、文献数の増加度合いに着目して可視化を行った。俯瞰図を縦横0.5×0.5のメッシュに分割して、メッシュ毎に成長曲線(ロジスティック関数)フィッティングを行い、立ち上がり年を求めた。領域ごとに立ち上がり年に基づいて可視化を行った結果を図20に示す。立ち上がり年が新しい(暖色系で表示)と今後の文献数の増加が予測され、立ち上がり年が古い(寒色系で表示)と増加ピークを過ぎたことを示している。ここでは論文バルクデータの収録状況を鑑みて2019年までのデータを用いて

可視化を行った。ファイトケミカル全般で文献数が増加傾向にあるため、全域的に今後の成長が見込まれる(立ち上がり年が2015年以降)結果となった。

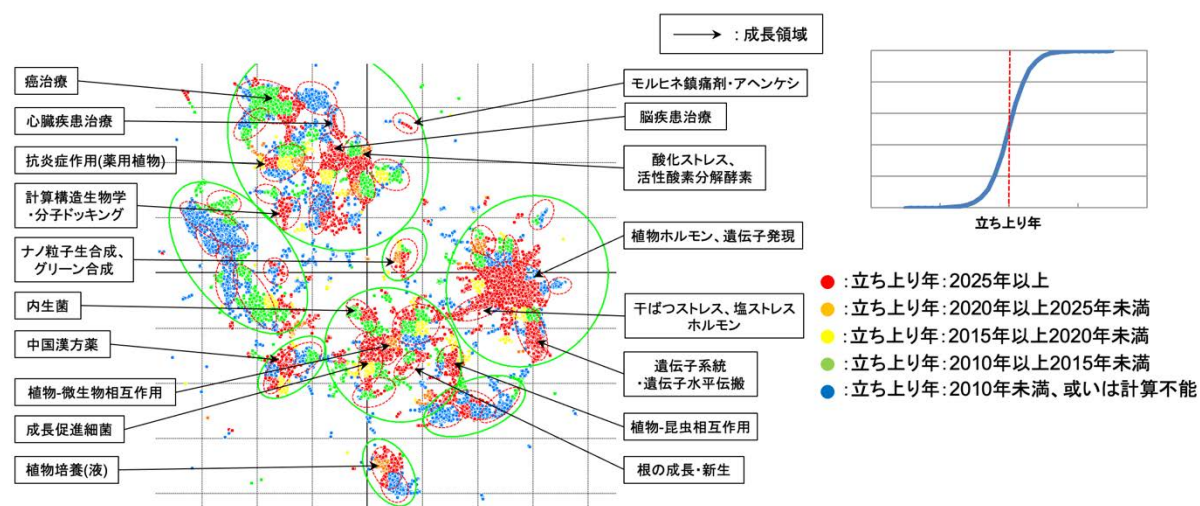


図 20 ファイトケミカル 論文+特許 成長曲線フィッティングによる立ち上がり年評価 絶対件数：全体

次に、論文・特許の文献数を全体平均で除して相対化することで、2020年の論文データも分析に使用して図20と同様の可視化を行った。単に文献数が増加しているだけでなく、全体平均以上に文献数が増加している領域を求めた結果を図21に示す。

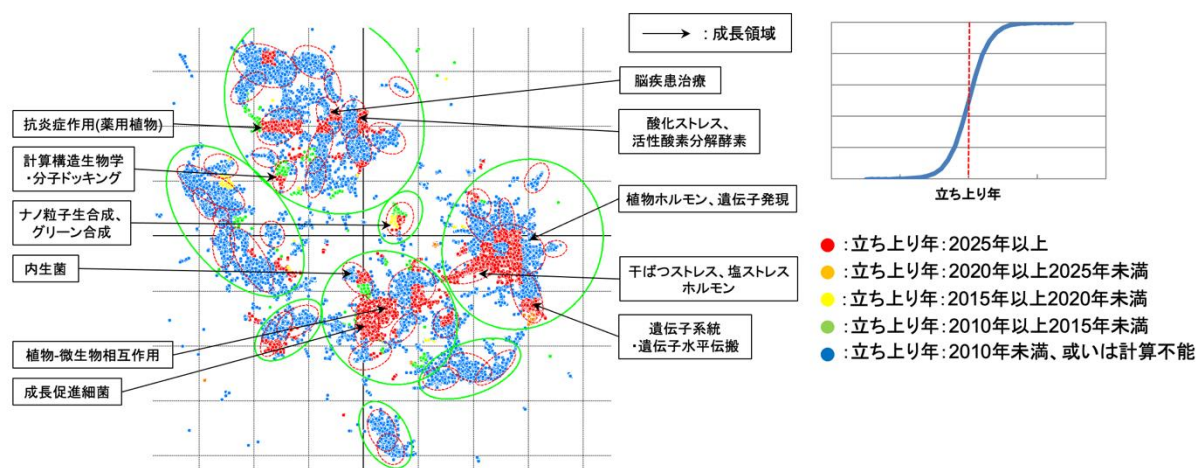


図 21 ファイトケミカル 論文+特許 成長曲線フィッティングによる立ち上がり年評価 相対件数：全体

ファイトケミカルの基礎的研究と思われる「植物ホルモン・遺伝子発現」「干ばつストレスホルモン」「遺伝子系統・遺伝子水平伝搬」「植物-微生物相互作用」「成長促進細菌」「内生菌」などで今後の成長性の可能性が高いと判定された。また、「脳疾患治療」や「酸化ストレス・活性酸素分解酵素」「抗炎症作用」などでも成長可能性が高くなっていた。

論文データのみで図21と同様の成長評価と可視化を行った結果を図22に示す。

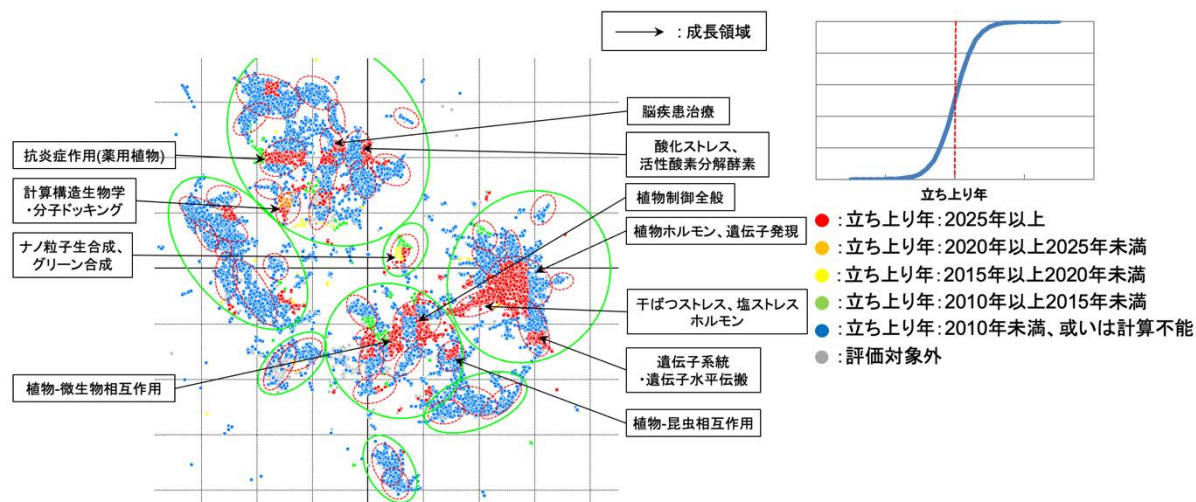


図 22 ファイトケミカル 論文 成長曲線フィッティングによる立ち上がり年評価 相対件数 : 全体

論文のみで評価を行った場合も、全体の文献に対して論文が占める割合が高い(85.6%)ため、大きな差は認められなかった。特許データのみで図21と同様の成長評価と可視化を行った結果を図23に示す。

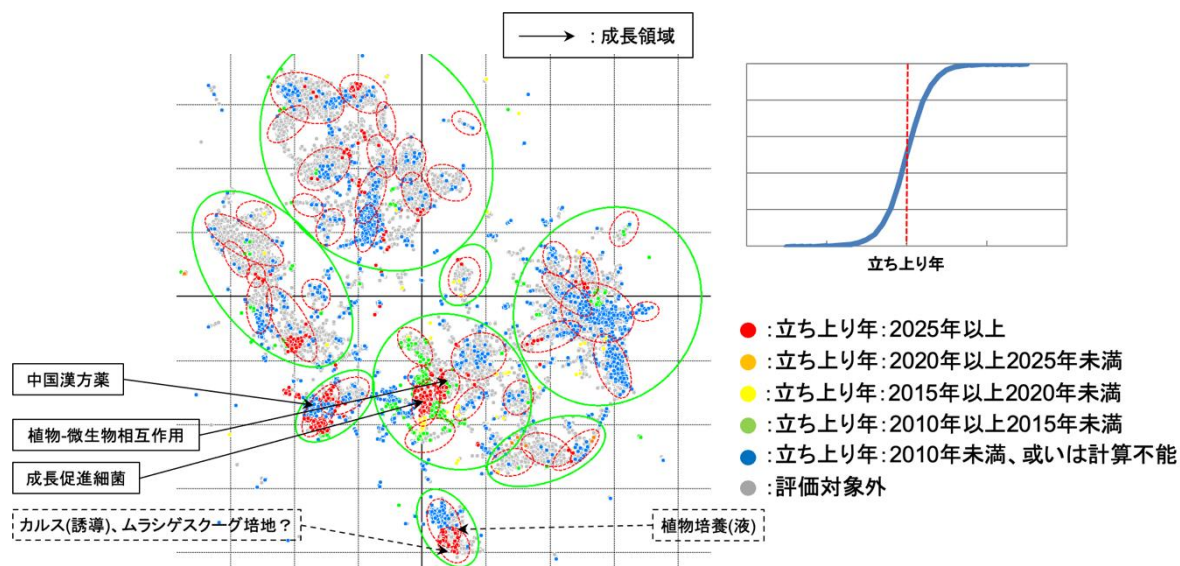


図 23 ファイトケミカル 特許 成長曲線フィッティングによる立ち上がり年評価 相対件数 : 全体

特許のみで評価を行うと、成長可能性が低いと判断される寒色系の領域が増加した。「成長促進細菌」「中国漢方薬」「植物-微生物相互作用」など、直近において特許出願数が増加している領域において、引き続き成長可能性があると判定された。なお、「カルス、ムラシゲスケーグ培地」の領域が成長判定と見えているが、実際には植物培養(液)の周辺部であると思われる。

ここまでエマージング領域を論文・特許数に基づいて可視化してきたが、研究に参入して論文を出版する研究機関が増加する領域においては、成長しているエマージング領域と判断することができる。そこで論文を出版した研究機関数の年次増加割合を指標として評価を行った。参入する研究

機関数が多い領域を求めた結果を図24に示す。なお、論文の研究機関名の同定には、WoSのorganization1を用いた。

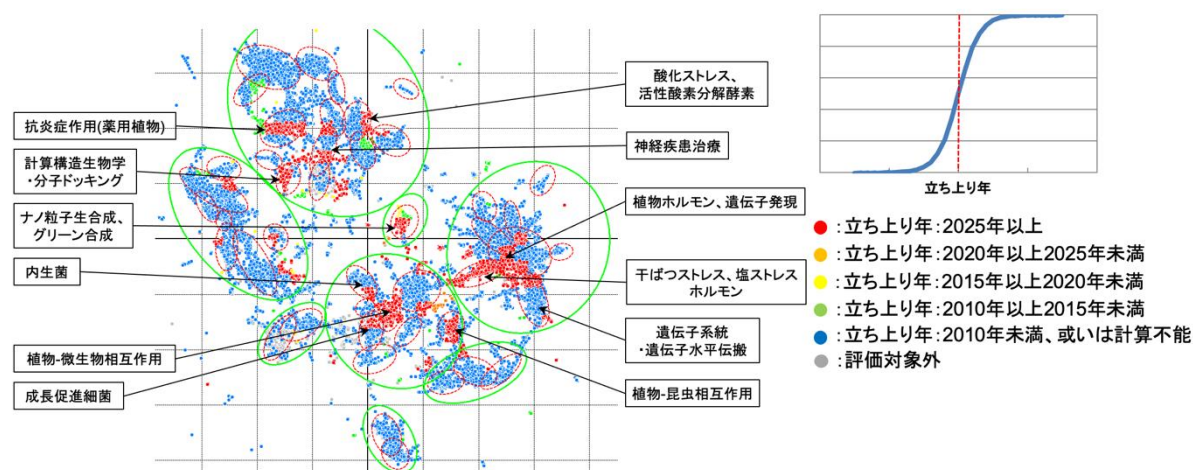


図 24 ファイトケミカル 論文 エマージングな研究領域の探索:成長曲線フィッティングによる立ち上がり年評価 相対件数：論文 研究機関

ユニークな研究機関数での評価は、論文件数による評価とあまり大差のない評価となった。論文件数を指標とした場合(図22)と比較して成長可能性領域と判定された領域が狭くなる傾向が見られた。「植物ホルモン、遺伝子発現」において成長可能性領域が狭くなった一方で「ナノ粒子生合成」「計算構造生物学・分子ドッキング」等の領域は今後成長可能性が高いと確認された。

次に特許に関してユニークな特許出願人の年次増減を用いて評価を行った。参入する出願人数が多い領域を求めた結果を図25に示す。

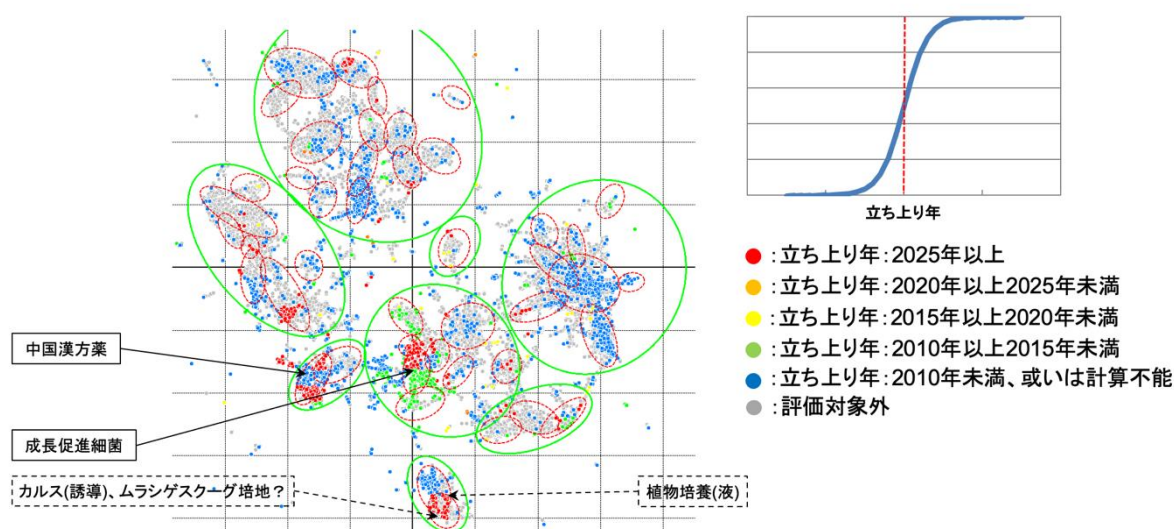


図 25 ファイトケミカル 特許 エマージングな研究領域の探索:成長曲線フィッティングによる立ち上がり年評価 相対件数：特許 出願人

ユニークな出願人数での評価は、特許件数による評価(図23)と大きな差が無い評価となったが、全体的に成長領域と判定された領域が狭くなる傾向が見られた。

次に論文を出版したユニークな論文著者数の年次増減をエマージング指標に用いて評価を行った。人的資源をより投入している領域を求めた結果を図24に示す。なお、論文著者の同定には、WoS付与のfull_name(roleはauthorに限定)を用いた。

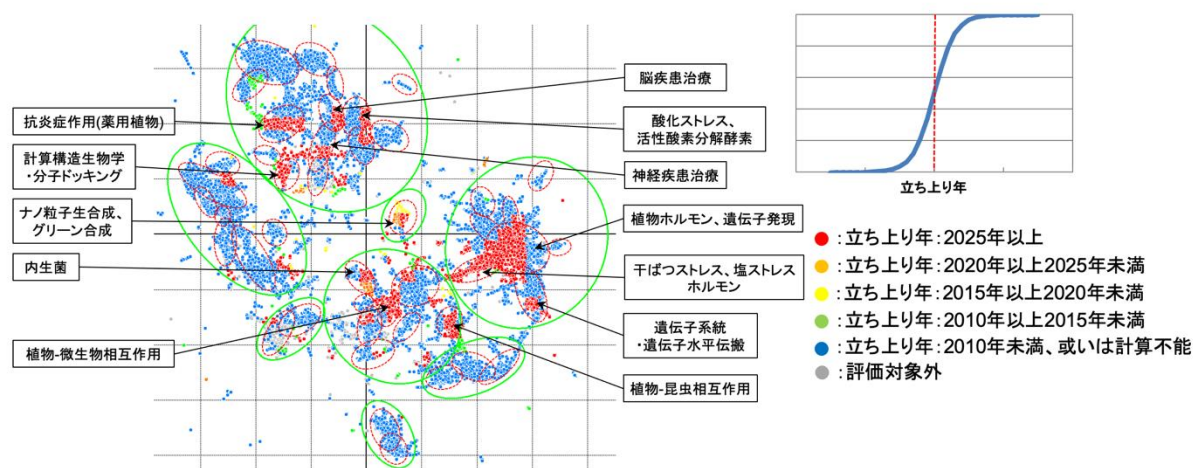


図 26 ファイトケミカル 論文 エマージングな研究領域の探索:成長曲線フィッティングによる立ち上り年評価 相対件数: 論文著者

論文のユニークな著者数を指標とした結果(図26)、論文数を指標とした結果(図22)と顕著な差は見られなかった。論文数、論文研究機関数、論文著者数のいずれの指標においても成長可能性があると判定された領域は確度が高いと推測できる。「植物ホルモン・遺伝子発現」「植物-微生物相互作用」「抗炎症作用(薬用植物)」「酸化ストレス・活性酸素分解酵素」などの領域が、論文数・新規参入する研究機関や研究者が増加しており、研究が活発になっているエマージング領域と考えられる。

3.5 国・地域別、機関別集計

ファイトケミカルに関する論文について、国・地域別の出版件数を図27(A)に、上位6カ国の論文数の推移を図28(A)に示す。European Patent Organisationに属する38カ国をまとめて欧州として集計した結果を図27(B)および図28(B)に示す。

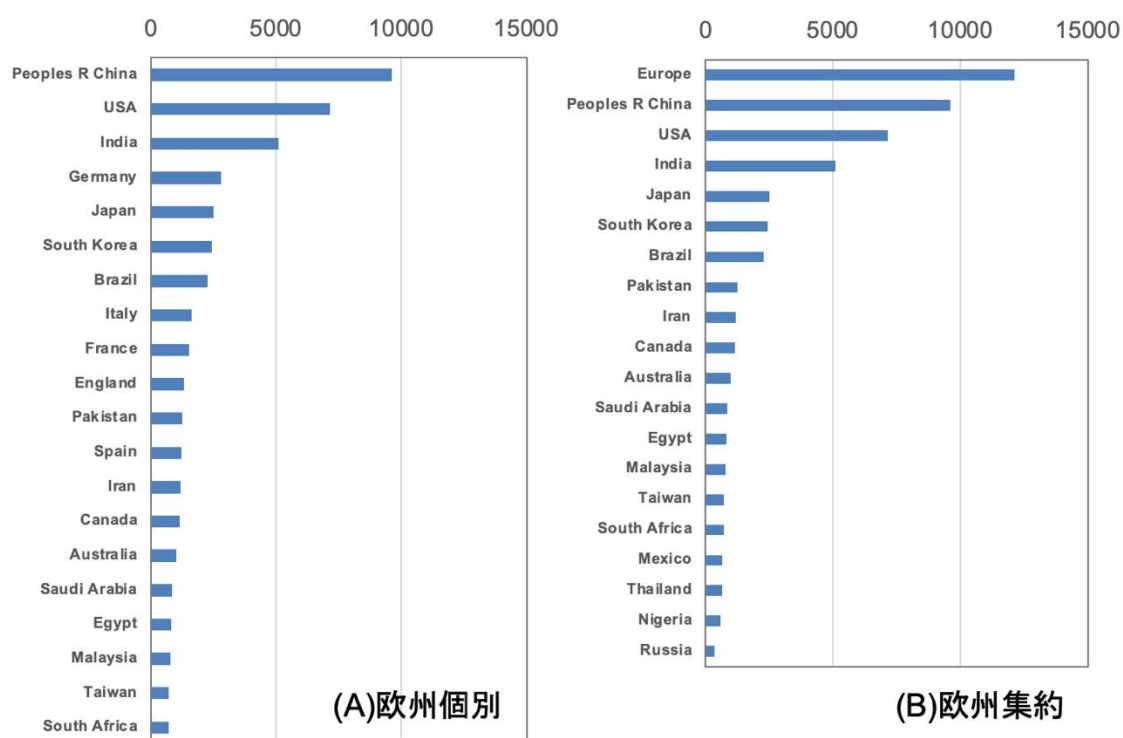


図 27 ファイトケミカル 論文 国・地域別件数:(A)欧州各国を個別に集計 (B)欧州を合算して集計

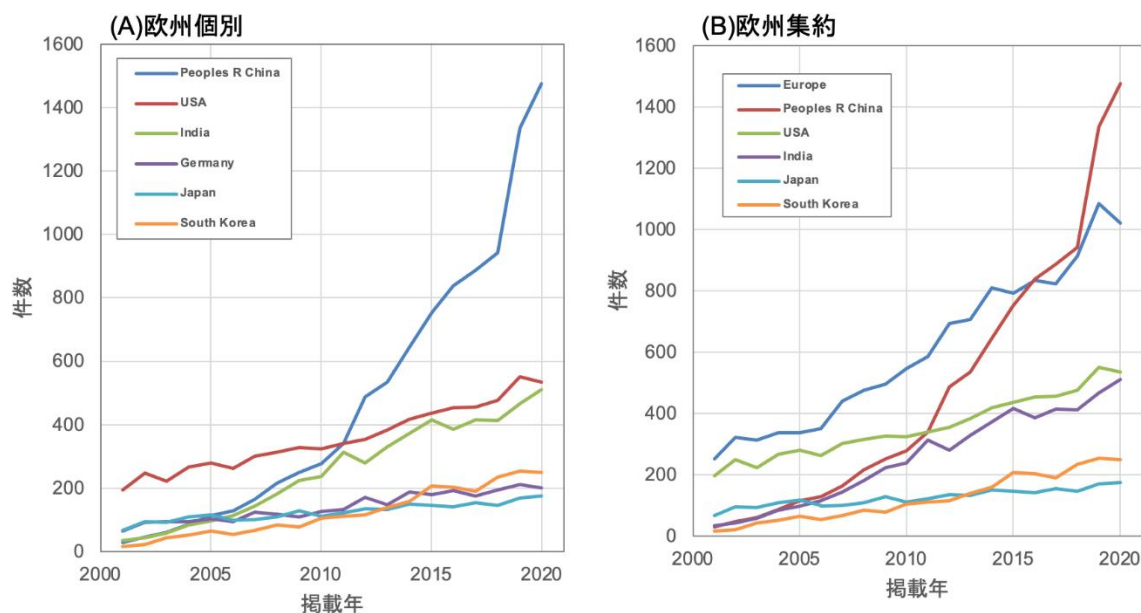


図 28 ファイトケミカル 論文 国・地域別件数推移:(A)欧州各国を個別に集計 (B)欧州を合算して集計

同様にファイトケミカルに関する特許について、国・地域別の特許出願件数を図29(A)に、上位6カ国の出願数の推移を図30(A)に示す。European Patent Organisationに属する38カ国をまとめて欧州として集計した結果を図29(B)および図30(B)に示す。なお、特許の国籍は発行国でなく優先権

主張先国とし、EPOに直接出願した特許などはEPのままとし、特定の国に加算せずにEPとして集計した。

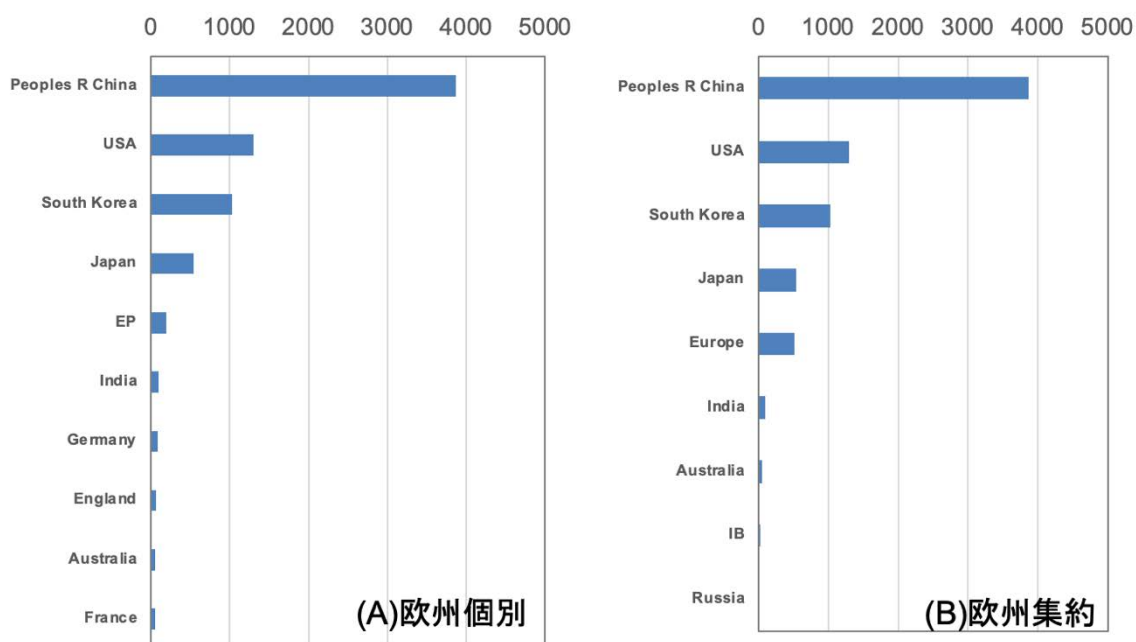


図 29 ファイトケミカル 特許 国・地域別件数:(A)欧州各国を個別に集計 (B)欧州を合算して集計

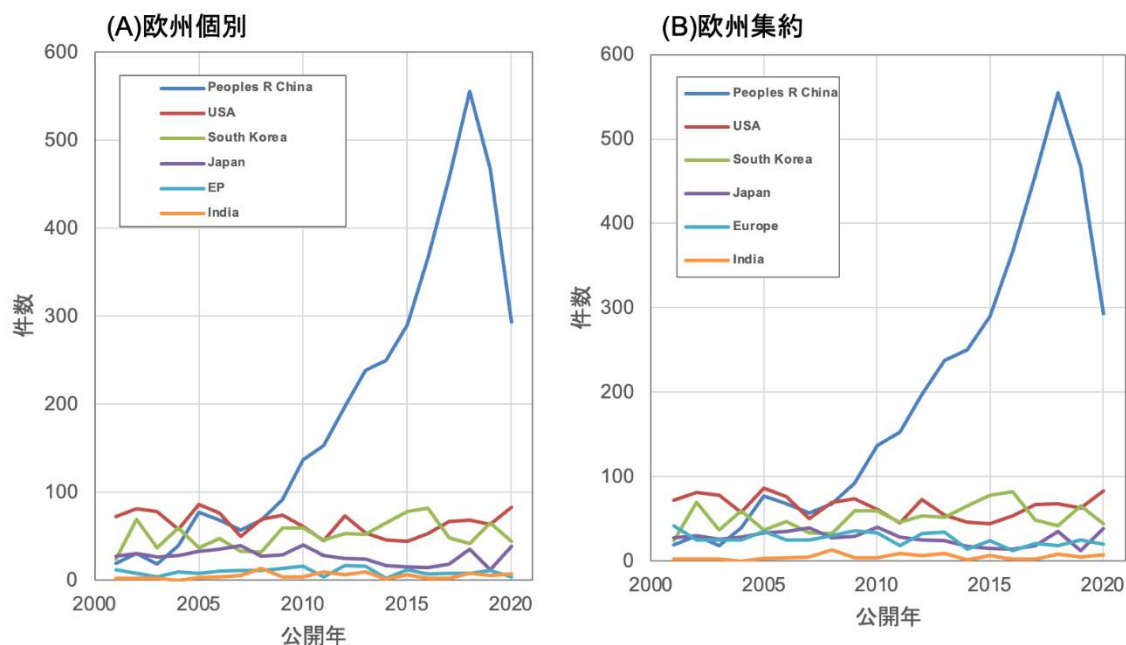


図 30 ファイトケミカル 特許 国・地域別件数推移:(A)欧州各国を個別に集計 (B)欧州を合算して集計

全体・論文のみ・特許のみのいずれの集計においても米国と中国が上位2国を占めている。3位は全体・論文のみ集計ではインド、特許のみ集計では韓国となる。4位は全体では韓国、論文のみ

集計ではドイツ、特許のみ集計では日本となり、これらの国がファイトケミカルにおける主要研究開発国と見られる。

ファイトケミカルで論文数・特許数の多かった中国・米国・欧州・インド・韓国・日本について、国ごとに論文出版・特許出願が多い領域を明らかにするため、論文44,712件及び特許7,486件からなる技術俯瞰図においてヒートマップによる可視化を行った。結果を図31に示す。

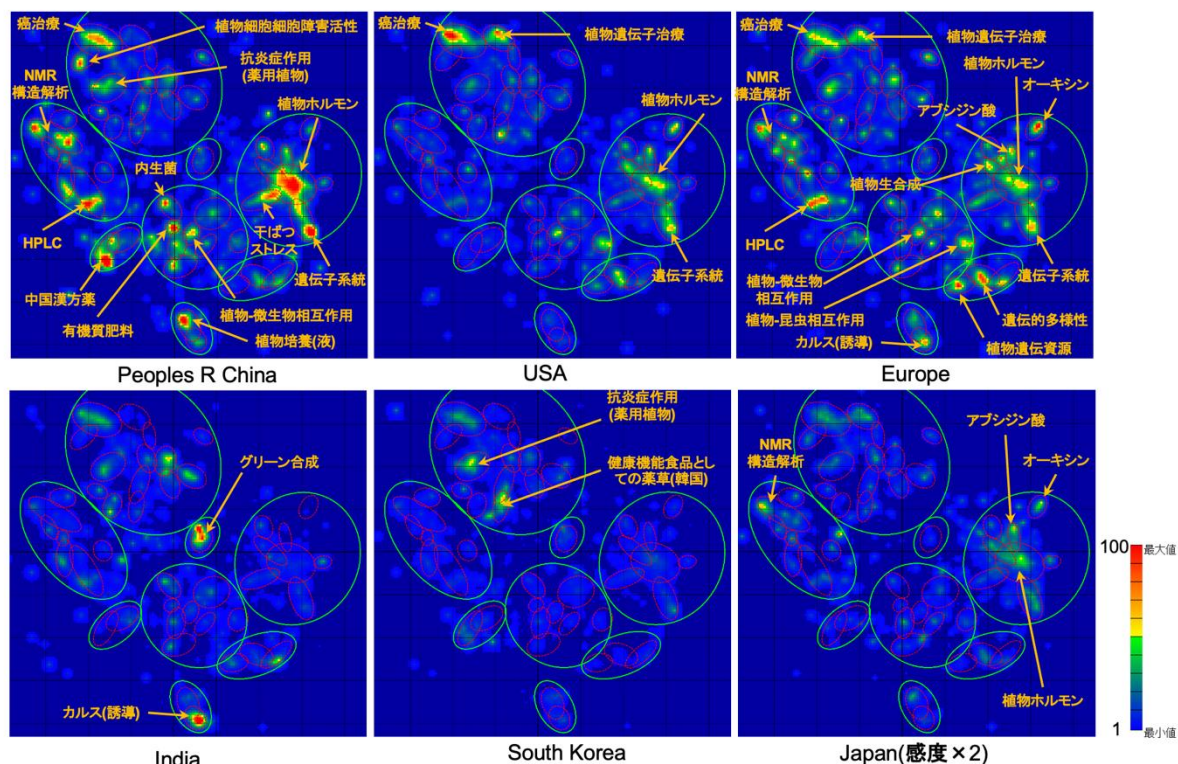


図 31 ファイトケミカル 論文+特許 トピックモデル計算の多様体学習による2次元可視化 :国・地域ごと

中国はファイトケミカルのほぼ全域で活動があり、特に「植物ホルモン」「遺伝子系統」「中国漢方薬」「植物培養液」「HPLC」などに集中している。米国は「癌治療」「植物遺伝子治療」「植物ホルモン」などに領域が限定されている。欧州では「遺伝子多様性」「植物遺伝資源」「オーキシシン」「HPLC」など米国以上に集中している領域が多い。インドは上位国では唯一「グリーン合成」に集中している。韓国は健康機能食品としての「薬草」や「抗炎症作用」など【植物による医薬・治療】に集中している。日本は「NMR構造解析」「アブシジン酸」「オーキシシン」「植物ホルモン」に集中している。

同様に国ごとに論文出版が多い領域を明らかにするため、論文のみからなる技術俯瞰図においてヒートマップによる可視化を行った。結果を図32に示す。

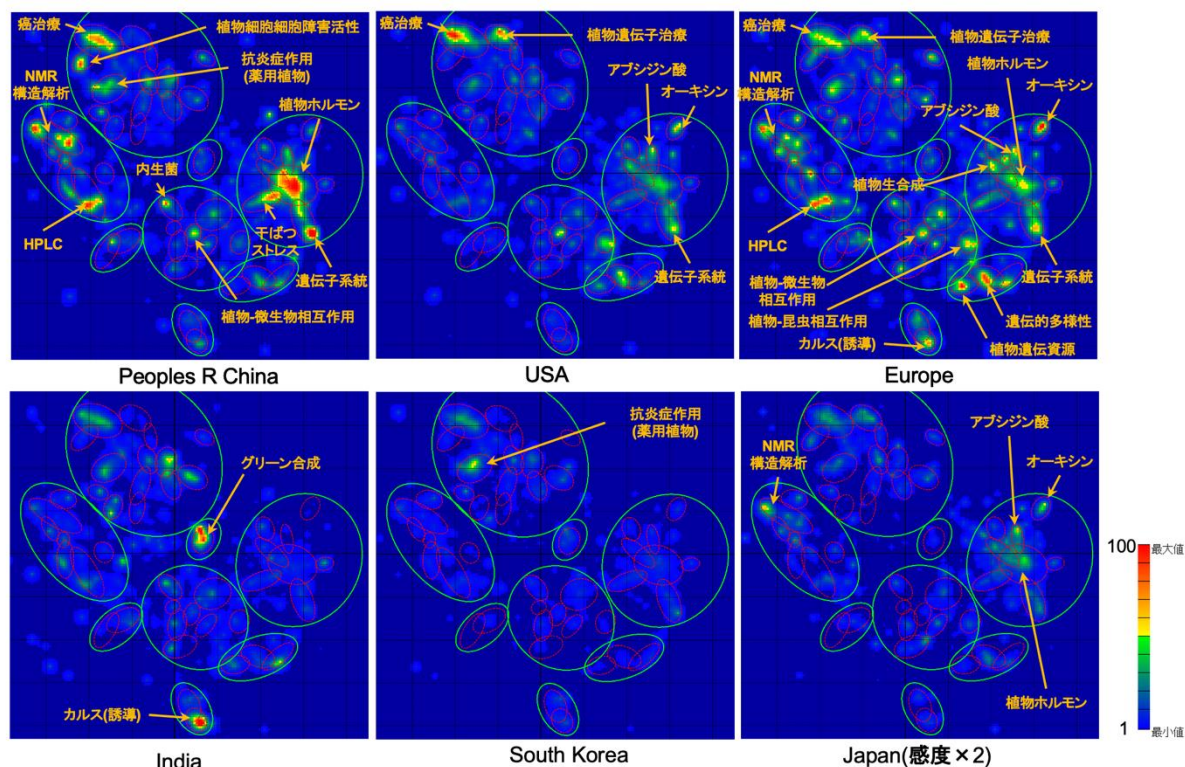


図 32 ファイトケミカル 論文 トピックモデル計算の多様体学習による2次元可視化 :国・地域ごと

中国は全体ではほぼ全域で集中領域が存在したが、論文のみで集計した場合は「植物ホルモンの作用」や「癌治療」で論文が出版されている。米国は論文の寄与が大きく、全体との差は小さい。欧州は特許出願件数が低いために全体集計と顕著な差がない。インドも、論文中心で全体との差異はほぼなく「グリーン合成」と「カルス(誘導)」に集中している。韓国では「抗炎症作用」などに領域が限定されている。日本も論文が主体で「NMR構造解析」等に集中している。

同様に国ごとに特許出願が多い領域を明らかにするため、特許のみからなる技術俯瞰図においてヒートマップによる可視化を行った。結果を図33に示す。

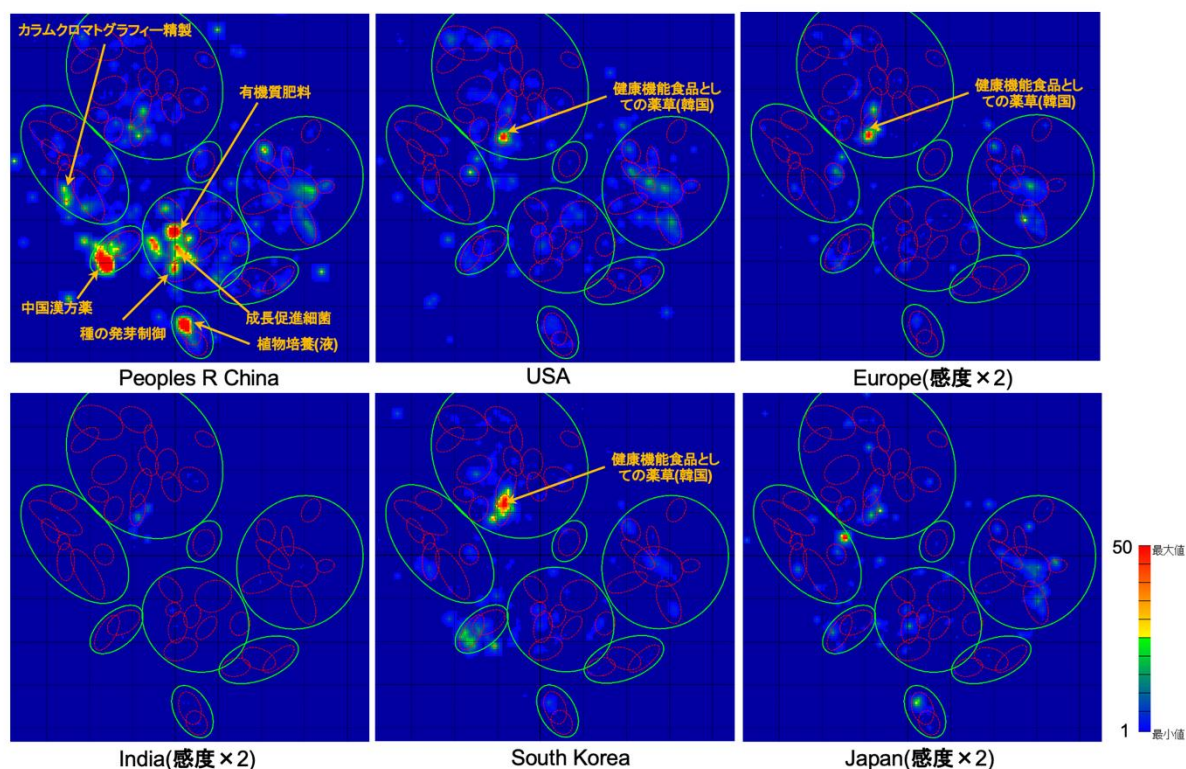


図 33 ファイトケミカル 特許 トピックモデル計算の多様体学習による2次元可視化 :国・地域ごと

特許は首位の中国が複数に出願している以外は、他の国では数カ所ではしか特許を出していない。中国の特許は「漢方薬」「有機質肥料」「植物培養液」等である。米国、欧州、韓国では「健康機能食品としての薬草」に特許が集中して、他の領域では少ない。日本も分子構造分析や植物培養にやや集中している程度であり、インドはほぼ特許を出願していない。

ファイトケミカルに関する論文について、出版した研究組織毎に出版数の集計を行った。なお、今回分析に用いたWeb of Scienceのバルクデータでは収録状況の影響により2020年において約1割程度の欠損が確認されている。論文出版数の集計結果を図34に、論文出版数の推移を図35に示す。

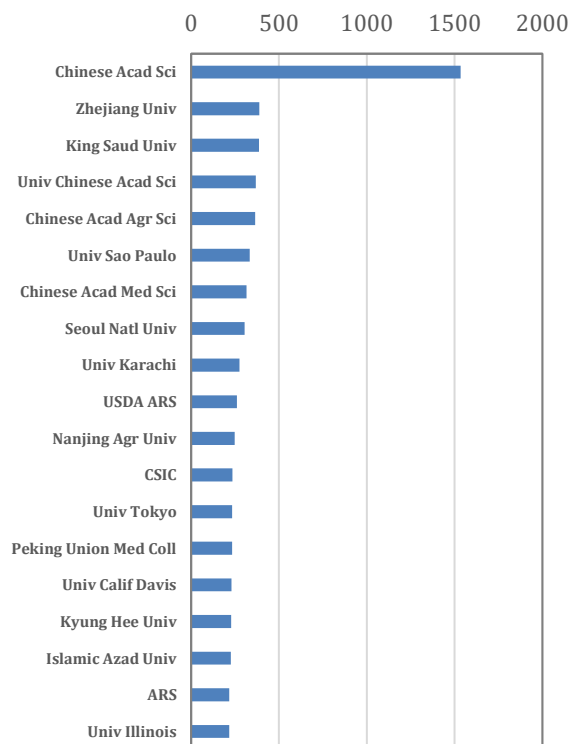


図 34 ファイトケミカル 論文 研究機関別件数

最も論文出版数の多い研究組織は中国の中国科学院となった。次いで中国の浙江大学、サウジアラビアのサウード国王大学、中国科学技術大学、中国農業科学院と中国が続く。米国の最高位は10位の米国農業研究事業団(USDA ARS)、日本の最高位は東京大学で13位となっていた。

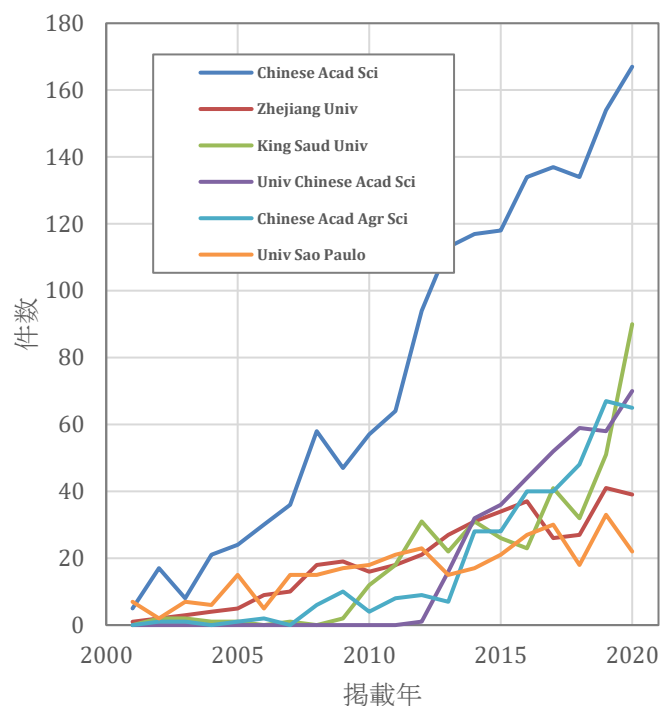


図 35 ファイトケミカル 論文 研究機関別件数推移

全ての上位機関が増加傾向のなか、首位の中国科学院の増加が著しい。3位のサウード国立大学も直近の増加が突出している。

技術俯瞰図において13の研究機関毎に論文数を示すヒートマップによる可視化を行った。結果を図36～図38に示す。

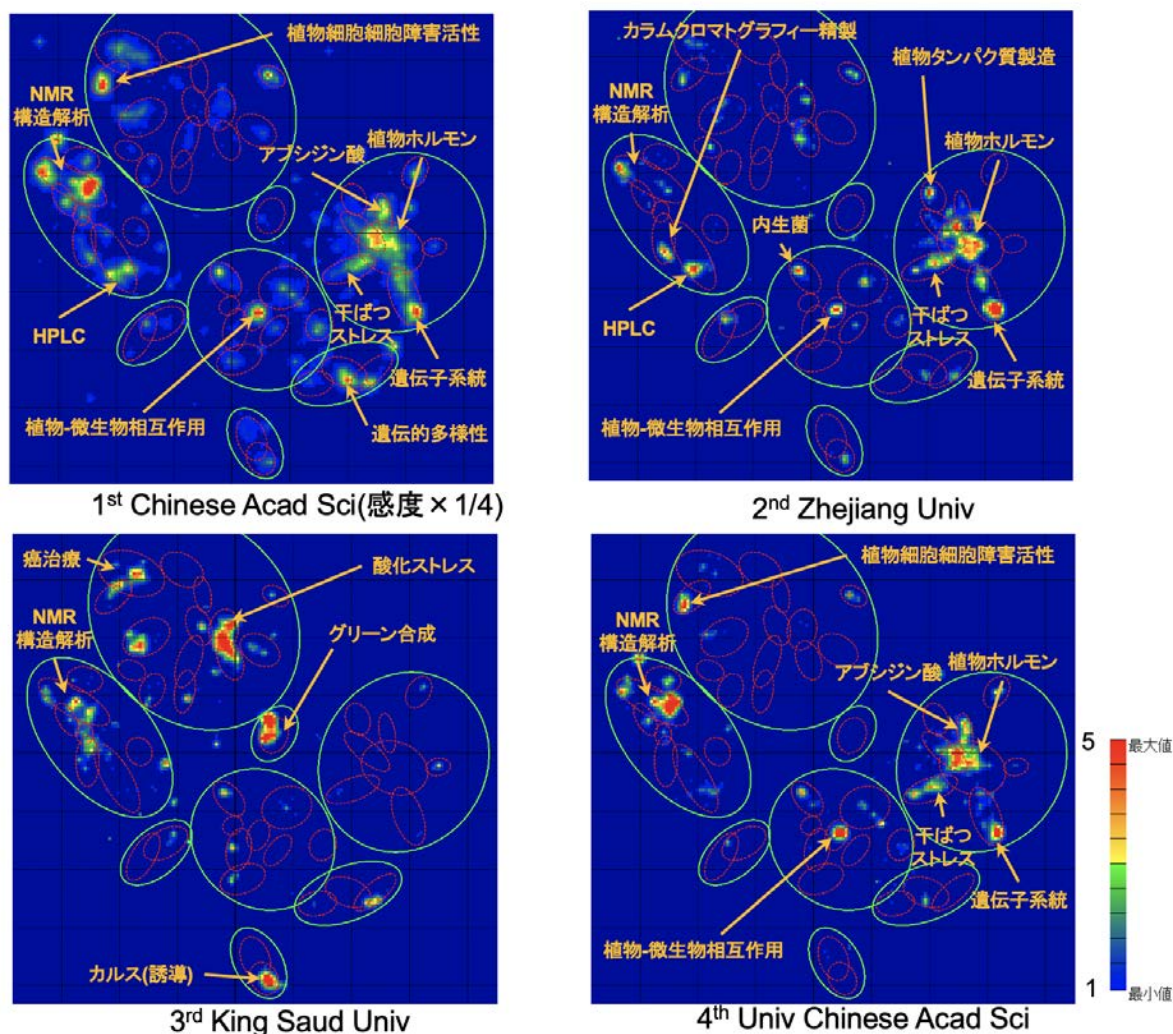


図 36 ファイトケミカル 論文 トピックモデル計算の多様体学習による2次元可視化 :研究機関ごと 論文数1位~4位

首位の中国科学院の件数が突出しているためヒートマップの感度を1/4にしている。「NMR構造解析」「植物-微生物相互作用」「遺伝子系統」「植物細胞細胞障害活性」等に集中している。2位の浙江大学は「遺伝子系統」「植物ホルモン」「HPLC」「植物-微生物相互作用」などに集中している。3位のサウジアラビアのサウード大学は「酸化ストレス」「グリーン合成」「カルス」に集中している。4位の中国科学技術大学は「NMR構造解析」「植物-微生物相互作用」「植物ホルモン」等に集中している。

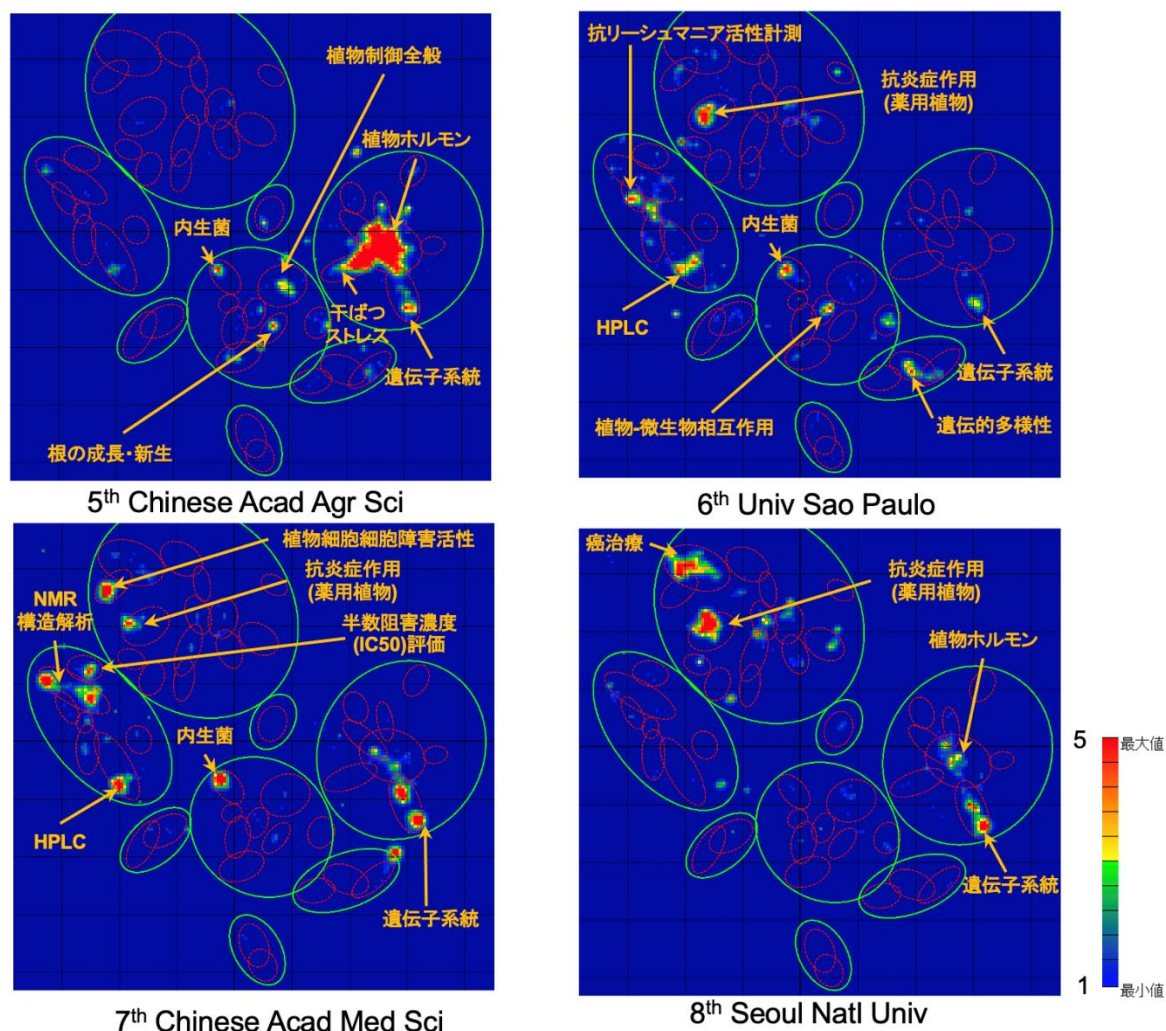


図 37 ファイトケミカル 論文 トピックモデル計算の多様体学習による2次元可視化 :研究機関ごと 論文数5位~8位

5位の中国農業科学院はほぼ「植物ホルモン」に特化している。6位のブラジル サン・パウロ大学は「抗炎症作用」「内生菌」「遺伝子多様性」などに集中している。7位の中国医学科学院は「遺伝子系統」「内生菌」「植物細胞細胞障害活性」などに集中している。8位のソウル大学は「癌治療」「抗炎症作用」「遺伝子系統」に集中している。

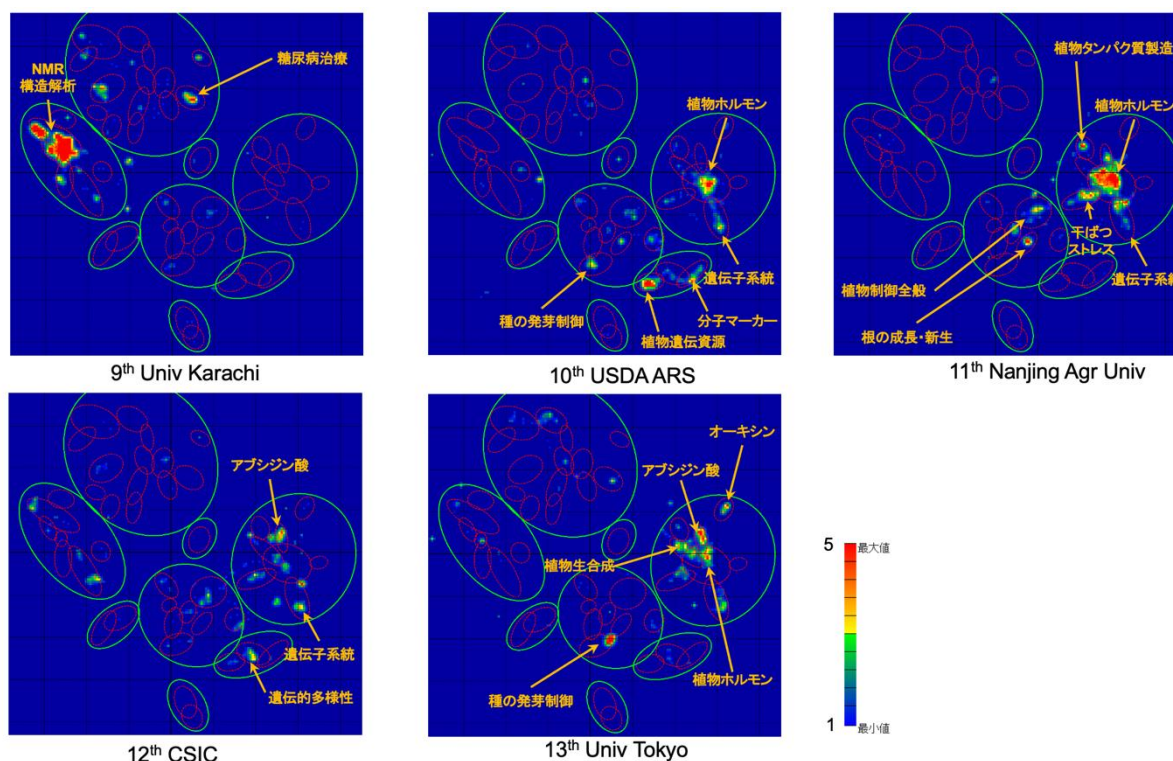


図 38 ファイトケミカル 論文 トピックモデル計算の多様体学習による2次元可視化 :研究機関ごと 論文数9位~12位

9位のパキスタン カラチ大学は「NMR構造解析」「糖尿病治療」に、10位の米国農業研究事業団 (USDA ARS)は「植物ホルモン」や「植物遺伝資源」などに、11位の南京農業大学はほぼ「植物ホルモン」に、12位のスペイン 科学研究高等評議会(CSIC)は「アブシジン酸」「遺伝子系統」などに13位の東京大学は「アブシジン酸」「オーキシシン」など植物ホルモンの集中している。

ファイトケミカルに関する世界の特許について、出願人毎に特許出願数の集計を行った。特許はファミリー重複削除を行っているため、登録年はファミリー内の登録特許中の最古年を採用した。特許出願数の集計結果を図39に、特許出願数の推移を図40に示す。

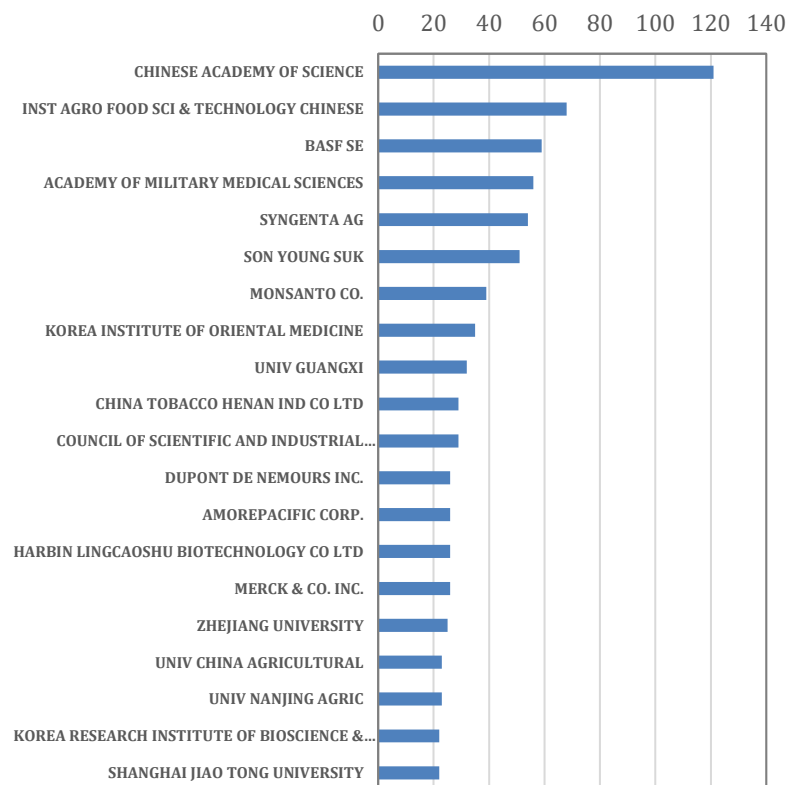


図 39 ファイトケミカル 特許 出願人別件数

出願人別特許出願件数の首位は中国科学院となった。2位は中国食品科学技術研究所(中国農業科学院内の組織)、3位はBASF、4位に中国 軍事医学研究院、5位にスイスのシンジェンタなど上位は中国中心のなか、欧州の企業も参入している。

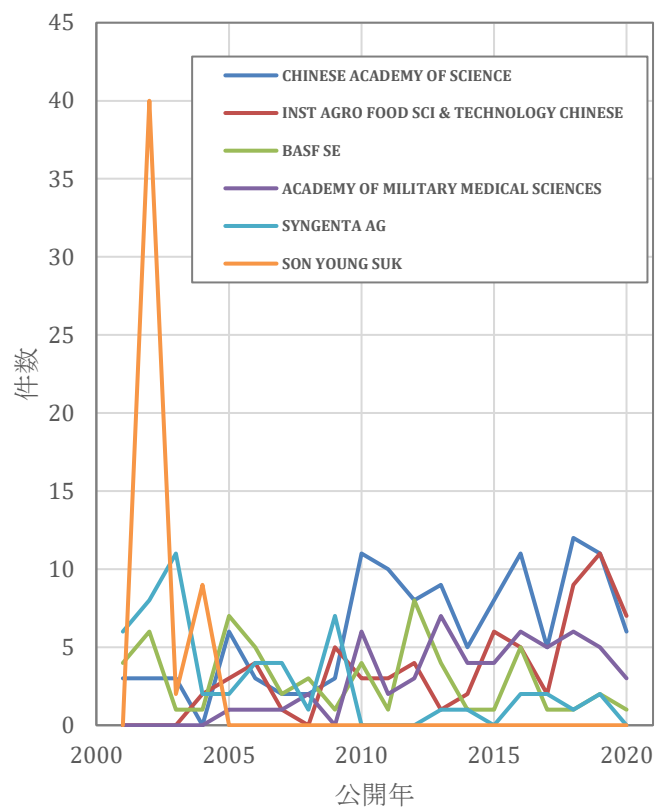


図 40 ファイトケミカル 特許 出願人別件数推移

中国科学院などの中国系の出願人がやや増加傾向にあるものの、明確な増加傾向は見当たらない。6位のSON YOUNG SUK氏は漢方薬関連のみに出願している個人の発明家と思われるため、考察の対象外とする。

技術俯瞰図において、12の機関毎に特許分布の可視化を行った。結果を図41～図43に示す。

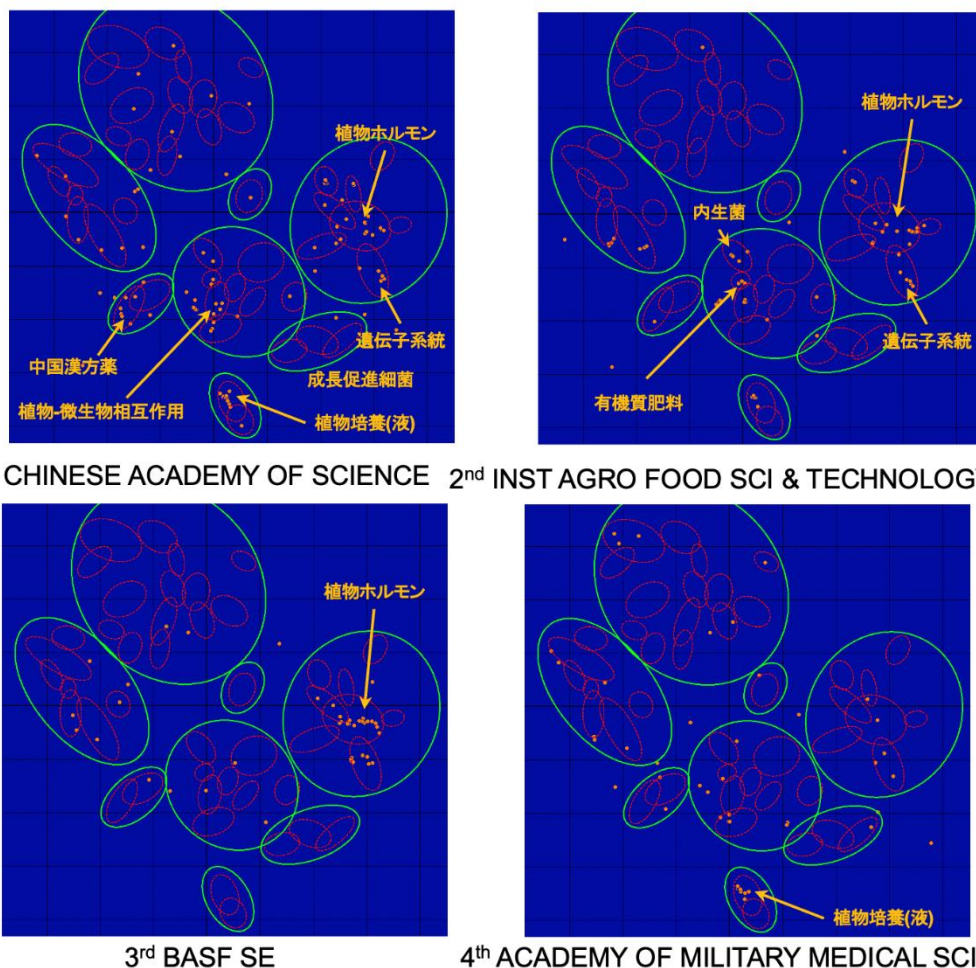


図 41 ファイトケミカル 特許 トピックモデル計算の多様体学習による2次元可視化 :研究機関ごと 特許出願数1位~4位

首位の中国科学院は「植物ホルモン」や「植物-微生物相互作用」「中国漢方薬」等に出願している。2位の中国食品科学技術研究所(中国農業科学院内の組織)も「植物ホルモン」「遺伝子系統」「有機質肥料」など基礎的な領域に出願している。3位のBASFも「植物ホルモン」に出願がある。4位の中国 軍事医学研究院は「植物培養液」に多い。特許全体の分布では「健康機能食品としての薬草」等の基礎研究ではない特許が多かった。しかし、出願人別上位で見ると、基礎的研究がなされていることが確認できた。

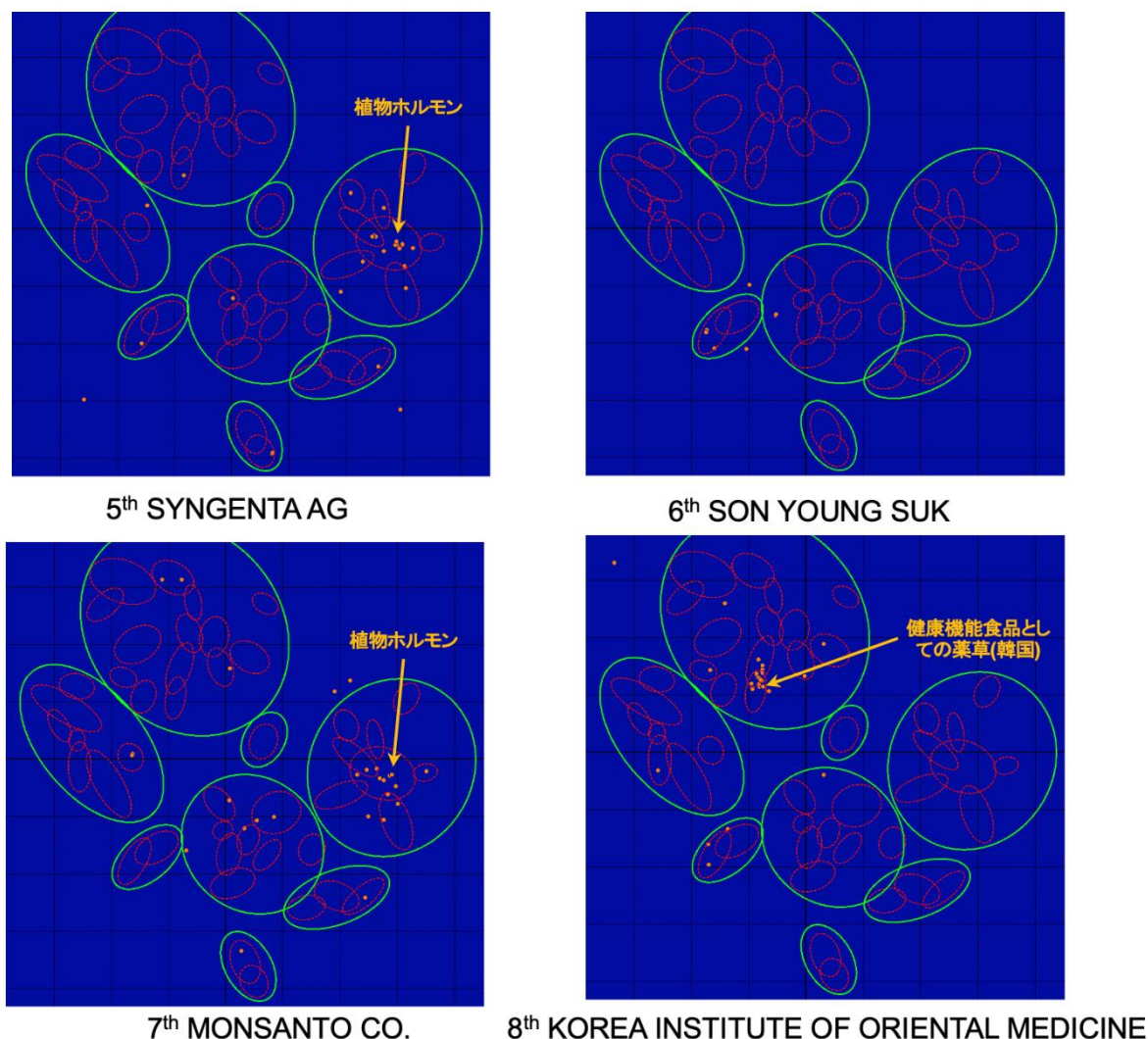


図 42 ファイトケミカル 特許 トピックモデル計算の多様体学習による2次元可視化 :研究機関ごと 特許出願数5位～8位

5位のスイス シンジェンタは「植物ホルモン」に集中している。6位の韓国SON YOUNG SUK氏は全て本人のみの単願で明らかな個人発明家であり、内容も「漢方薬」となっていた。7位のモンサント(現在はバイエルにより買収)も「植物ホルモン」に集中している。8位の韓国東洋医薬研究所は「健康機能食品」としての薬草に集中している。

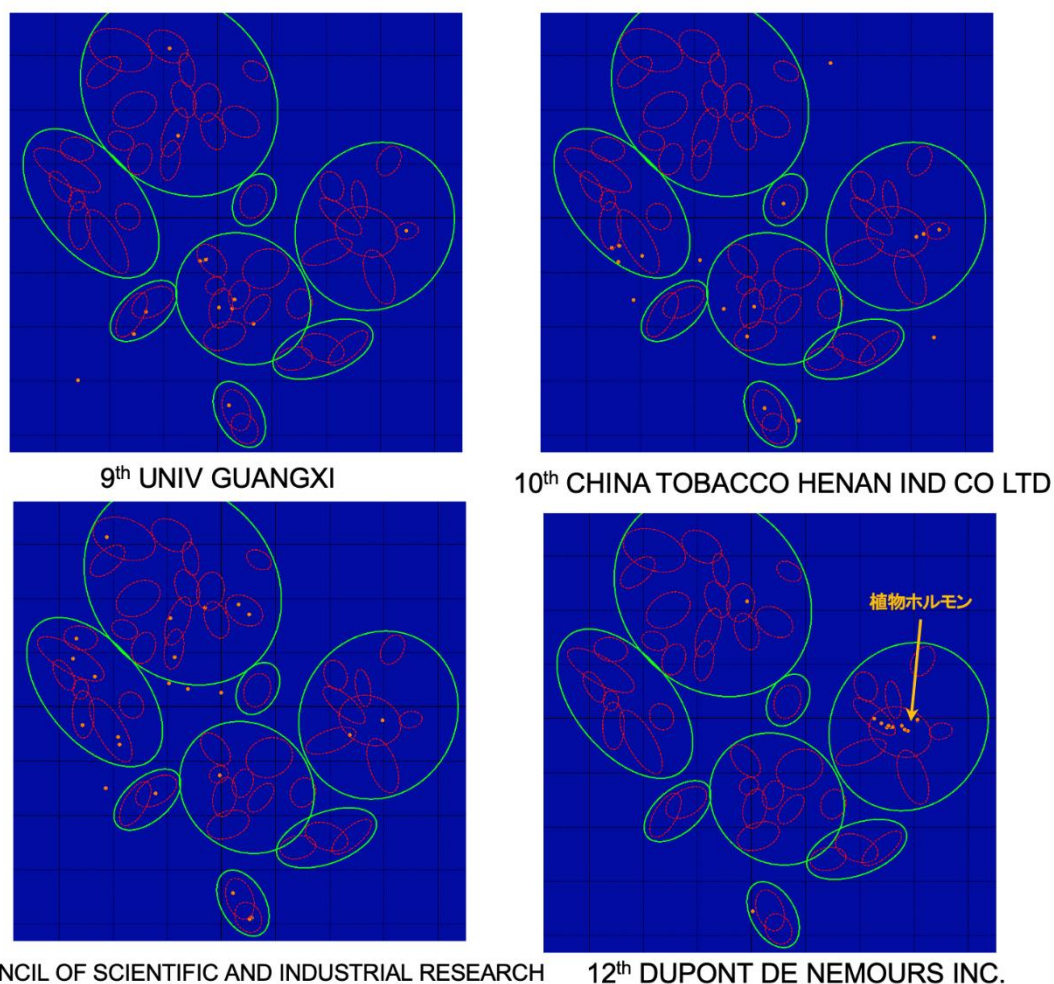


図 43 ファイトケミカル 特許 トピックモデル計算の多様体学習による2次元可視化 :研究機関ごと 特許出願数9位~12位

9位の中国 広西大学、10位の河南中煙工業、11位のインド科学産業研究委員会は特に集中した領域が見られない。12位のデュポンは「植物ホルモン」に注力している。

4 まとめ

ファイトケミカルに関連する論文・特許を抽出し、タイトル・アブストラクト等のテキストデータを用いたクラスタリング分析を行うことで、学術研究および技術開発の動向を二次元的に可視化した俯瞰図を作成した。

論文44,712件と特許7,486件から作成した俯瞰図から、【植物による医薬・治療】【植物ホルモン・遺伝子発現】【生合成制御】【分子構造分析・同定】の4つの大きな領域と【グリーン製造】【遺伝資源】【植物培養】【薬草・漢方薬】のやや小さな4つの領域の合計7領域に分かれている様子が確認された。

論文数・特許出願数は一定の割合で増加しており、2011年以降から「植物ホルモン」「内生菌」「成長促進細菌」などの研究が始まっている。また応用先としての「癌治療」「抗炎症作用」「糖尿病治療」や、補助的な計測技術である「NMR構造解析」「HPLC」が増加している。2016年以降では基

礎技術に近い「植物ホルモン」「植物-微生物相互作用」「成長促進細菌」「グリーン合成(金属ナノ粒子)」の増加が著しい。

論文数と、論文を出版した研究機関・研究者数の大きな増加が予測されるエマージングな研究領域として「植物ホルモン・遺伝子発現」「干ばつストレスホルモン」「遺伝子系統・遺伝子水平伝搬」「植物-微生物相互作用」「成長促進細菌」「内生菌」「脳疾患治療」「酸化ストレス・活性酸素分解酵素」「抗炎症作用」「計算構造生物学・分子ドッキング」「ナノ粒子生合成、グリーン合成」が抽出された。

特許数と特許出願人数の大きな増加が予測されるエマージングな研究領域として「成長促進細菌」「中国漢方薬」が抽出された。

今回の調査ではCRDS(2021)で記載されたファイトケミカル関連の学術研究領域の整理を行うことができたが、特許においては健康機能性食品や漢方薬等の天然植物由来成分を活用した技術が多く見られた。「成長促進細菌」や「植物-微生物相互作用」等の学術研究とリンクする技術開発が進んでいる分野が確認できたものの、2020年段階では基礎的な理解を深める段階であり、将来的に特許数が増加していくものと考えられる。

謝辞

CRDS桑原 明日香フェローに文献抽出の条件設定に関する議論で協力頂いた。

本研究は文部科学省共進化実現プログラム池内プロジェクト「科学技術・イノベーション政策の経済社会効果分析の政策形成プロセスへの実装」の助成を受けて実施された。

参考文献

黒田 昌裕, 池内 健太, 原 泰史

2016.「科学技術イノベーション政策における政策オプションの作成 -政策シミュレーターの構築- (モデル構築編)」, 政策研究大学院大学 科学技術イノベーション政策研究センター SciREX Working Paper 2016-#01, <http://doi.org/10.24545/00001570>

CRDS調査報告書

2016.「科学技術イノベーション政策の科学における政策オプションの作成～ICT分野の政策オプション作成プロセス～」, CRDS-FY2015-RR-07、2016年3月

CRDS戦略プロポーザル

2021.「ファイトケミカル生成原理とその活用のための研究開発戦略 ～未利用植物資源から革新的価値を創出する学術基盤の創成～」, CRDS-FY2020-SP-03、2021年3月



SciREX Center



政策研究大学院大学
NATIONAL GRADUATE INSTITUTE
FOR POLICY STUDIES

科学技術イノベーション政策研究センター

Science for RE-Designing Science, Technology and Innovation Policy Center (SciREX Center)

〒106-8677 東京都港区六本木 7-22-1 / Tel 03-6439-6318/ Fax 03-6439-6260

7-22-1 Roppongi, Minato-Ku, Tokyo 106-8677 JAPAN

Tel +81-(0)3-6439-6318 / Fax +81-(0)3-6439-6260