

中国科学院国家科学图书馆青年人才领域前沿项目研究报告

## 基于专利引文的知识转移研究

课题负责人：王学昭

课题组成员：赵亚娟、刘锐、张静

中国科学院国家科学图书馆总馆

北京

2012年11月

## 摘 要

从基础研究到技术创新的知识转移过程研究一直以来颇受关注。专利引文分析可以利用专利来分析基础研究和应用技术的联系，为深入探索基础研究对技术创新的影响提供了一种方法。本项目提出了基于专利引文知识转移分析模型，探索技术领域的知识转移情况，了解技术创新与科学基础研究的联系。该模型避免了传统知识转移研究方法缺乏普适应、指标难以衡量等局限性，适用于所有技术领域的知识转移研究。主要研究成果包括：

### 1、建立知识转移分析模型

基于专利引文，利用共被引聚类分析方法，本项目建立了知识转移分析模型。通过知识转移分析模型，可以揭示研究领域内的技术热点、科学基础知识源，并且表征了二者之间的知识转移路径和强度，最终形成知识转移地图。该分析模型立足于专利文献公开信息，侧重于定量化手段揭示领域内的知识转移过程，具有较强的实用性和可操作性，在专利分析中具有一定的参考价值。

### 2、揭示了转基因水稻技术领域的知识转移

以转基因水稻技术领域为研究对象进行知识转移模型的实证分析，得到了较为合理且客观的该领域知识转移的分析结果，揭示了 8 个技术热点和 10 个科学基础知识源的联系。

### 3、形成知识转移分析工具

本项目形成了一套知识转移分析工具，在该分析工具中对分析的指标和参数的设置均进行了详细说明，可为今后同类工作借鉴。

**关键词：** 专利引文；知识转移；共被引；专利家族

## Abstract

The research on knowledge-transfer from basic research to technology innovation is always focused on. Patent citations provide a key way of exploring the effect of basic research on technology innovation. Based on co-citation cluster analysis, we propose a knowledge-transfer analysis model for patents, with which the landscape of knowledge-transfer including route and strength between knowledge sources and technology hotspots can be mapped out.

By means of co-citation cluster analysis on cited articles, the scientific knowledge sources (articles clusters) of a certain technology field can be retrieved. Forward patent citations can reveal the technology hotspots of a certain technology field, which can be obtained through patent co-citation cluster analysis. The route and strength of knowledge-transfer is relative to the citation of the patents in patent clusters to the articles in article clusters. Ultimately, this knowledge-transfer analysis model would totally present the knowledge-transfer from scientific knowledge to technology in a certain technology field.

The model has been applied to patents from the field of transgenic rice. As a result of the analysis, 10 scientific knowledge sources and 8 technology hotspots emerge, and reasonable links between have been established, which clearly show how knowledge has been transferred.

Avoiding the limitations of traditional methods of knowledge-transfer, this model is a principle applicable to all fields of technology.

**Keywords:** patent citation; knowledge-transfer; co-citation; patent family

## 目录

一、引言 .....	1
1.1 研究背景.....	1
1.2 研究内容和意义.....	2
1.3 研究方法.....	3
二、研究模型和数据 .....	4
2.1 从科学技术知识到技术创新的知识转移模型.....	4
2.2 知识转移模型分析流程.....	4
2.3 数据获取.....	5
2.4 高被引论文或专利.....	7
2.5 共被引关系的建立.....	7
2.6 共被引聚类分析.....	8
2.7 知识转移表征.....	9
2.8 可视化.....	10
三、实证分析——转基因水稻 .....	11
3.1 转基因水稻简介.....	11
3.2 转基因水稻技术领域的科学基础知识源.....	12
3.3 转基因水稻技术领域的技术热点.....	14
3.4 转基因水稻技术领域的知识转移.....	16
四、结论 .....	17
4.1 研究成果.....	17
4.2 存在问题.....	18
4.3 未来研究方向.....	18
参考文献: .....	20

# 一、引言

## 1.1 研究背景

从基础研究到技术创新的知识转移是一直以来颇受关注的问题。如何度量科学基础研究和技术创新之间的关系，展现和解释从科学到技术的知识转移过程，了解技术创新对科学基础研究的依赖程度，是一个富于挑战性而又极具研究价值的问题。

对于科学基础研究到技术创新的知识转移，文献中记载的方法有案例研究法和费用—效益分析、计量经济学方法、专利引文分析等<sup>1-2</sup>。案例研究法能够追溯某一特定领域从基础科学知识到技术的进步，但缺乏普适性；费用—效益分析方法以分析个别项目的效益为出发点，计量经济学的方法则基于研究技术研发投入与经济增长的关系，但由于研发的产出难以衡量，所以二者也有一定局限性。

专利作为创新的重要产物，是基础研究和技术的共同体现，对联系科学知识和技术其中举足轻重的作用。专利对非专利文献引用（前引）可以展现技术创新对科学知识的依赖，专利的被引用（后引）体现了专利对技术创新的推动。专利引文分析利用专利连接基础研究和应用技术，可以揭示基础研究对技术创新的影响，发现知识转移的路径和强度。

用专利测度技术创新，用非专利文献引文测度科学基础研究，对专利引文中的非专利文献进行计量分析，反映出技术创新对科学基础研究的依赖程度，或者说反映出科学基础研究对技术创新的贡献大小<sup>3</sup>。Jaffe、Duguet 等人的研究表明专利引文与知识的获取和发送之间存在较高的一致性，专利引文可以作为知识转移的度量工具<sup>4</sup>。美国 CHI 研究公司的 Narin 开拓了专利引文分析的新思路，即通过对专利文献所引用的科学论文进行计量分析，揭示公共科学对技术创新的贡献程度<sup>5-6</sup>。Meyer 通过比较不同领域专利文献与期刊文献的引用数量关系，反映技术对科学基础研究的依赖程度<sup>7</sup>。陈超美尝试从使用复杂网络统计方法、网络可视化及引文分析的方法揭示科学与技术之间的知识转移关系[8]。近期，日本科学技术振兴机构（Japan Science and Technology Agency, JST）基于专利以及专利所引用的非专利文献（non-patent literature, NPL）信息研究了技术联系强度

(technology linkage)<sup>9</sup>。OECD 提出了平均每个专利所引用的非专利文献和施引专利的份额等指标研究专利的技术与科学关联性<sup>10</sup>。

共被引分析是研究专利引文分析的重要方法之一。同时被另一个元素所引用的两个元素为共被引关系。Henry Small 于 1973 年首先开始了论文的共被引研究<sup>11</sup>，之后共被引分析被广泛的应用于科技政策制定与评估中。Mogee 应用专利共被引聚类分析揭示了美国礼来公司 (Eli Lilly and Company) 专利的技术前沿<sup>12</sup>。Kuei-Kuei Lai 利用专利共被引分析建立新的专利分类体系代替 IPC (International Patent Classification, 国际专利分类号) 和 USPC 分类系统，用于专利分析和管管理<sup>13</sup>。王贤文基于专利共被引方法对世界 500 强中的工业企业的大型专利共被引网络分析<sup>14</sup>。

然而，专利共被引聚类分析在知识转移的研究中却较比罕见。

## 1.2 研究内容和意义

本研究在目前国内外研究的基础上进行深入研究，基于专利引文，采用共被引聚类分析的方法，形成知识转移分析模型和方法，直接且客观地揭示技术创新和基础研究的联系，了解技术领域的知识转移现状。

项目的最终目标是形成基于专利引文同被引的知识转移分析模型和方法，展现目标研究领域内技术热点，科学基础知识源，知识转移现状 (如图 1-1 所示)，了解技术领域中技术创新对科学基础研究的依赖程度，即科学基础知识对技术创新的支撑程度。

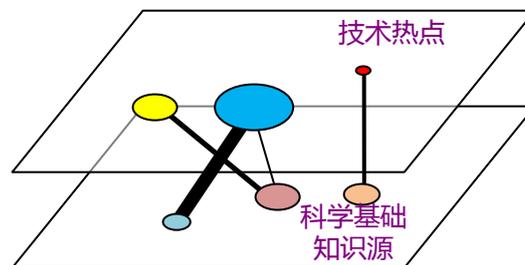


图 1-1 知识转移过程

### 1.3 研究方法

本项目针对研究过程中可能出现的问题，拟采取如下研究方法：

#### (1) 文献调研

通过各种信息渠道，追踪相关研究的最新动态，熟悉并掌握可利用的资源和工具。

#### (2) 模型设计的理论和方法

通过同被引聚类模型设计，采用定量分析的方法，从而避免了人的主观经验判断的不足，可以客观、准确地对专利进行分析。

#### (3) 实例验证

采用实例验证的方法来进行对本项目所提出的分析模型和方法的改进，并证实其有效性和实用性。

#### (4) 专家咨询

在实例分析的过程中和相关领域的研究人员保持深入的沟通，保证分析结果的准确性。

本项目研究方案如图 1-2 所示。

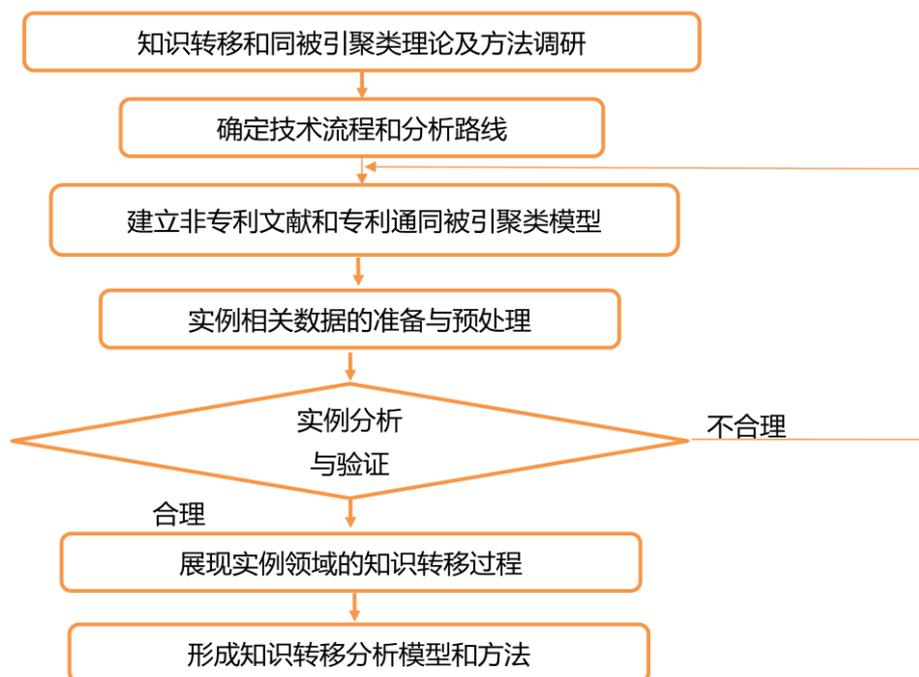


图 1-2 基于专利引文的知识转移研究分析方案设计

## 二、研究模型和数据

### 2.1 从科学技术知识到技术创新的知识转移模型

#### 1) 确定科学基础知识源和技术热点

专利的引文信息包括专利所引用的专利文献和非专利文献，同时某些数据库也会给出专利的被引用情况。专利对非专利文献引用（前引）可以展现技术创新对科学知识的依赖，专利的被引用（后引）体现了专利对技术创新的推动。该研究建立的知识转移模型通过对专利（如图 2-1 中的专利 A 和 B）所引用的非专利文献（如图 2-1 中的非专利文献 1-4）的共被引关系进行聚类，进行内容分析，确定主要科学基础知识源；通过专利（如图 2-1 中的专利 A 和 B）被后续专利（即：施引专利，如图 2-1 中的 I、II 和 III）所引用的情况进行聚类，形成技术分布结构，进行内容分析，确定技术热点。

#### 2) 知识转移表征

利用技术热点中专利对科学基础知识源中非专利文献引用关系确立知识转移路径，考虑引用频次和相应技术热点及科学基础知识源的大小的关系，建立主要科学基础知识源与技术热点之间的关联强度，了解研究知识转移的趋势。

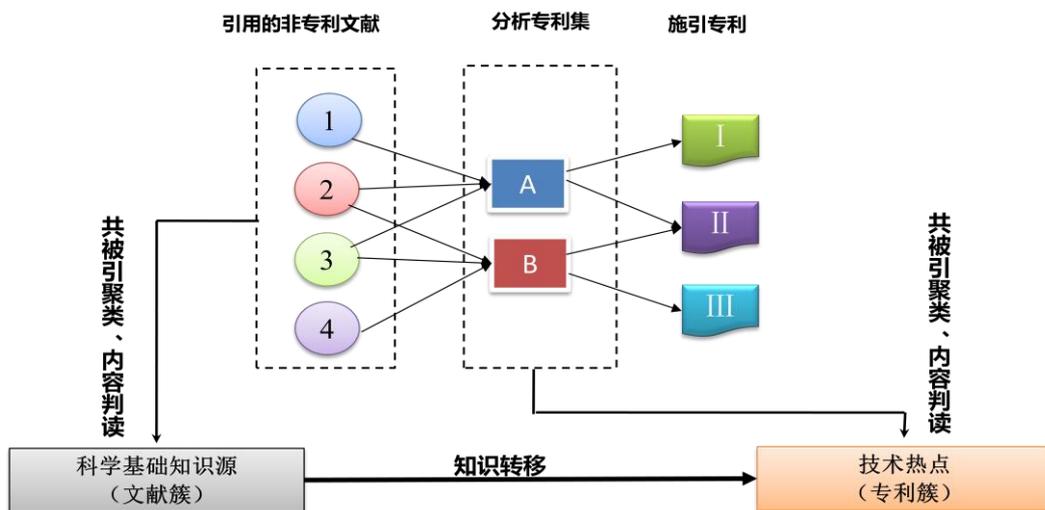


图 2-1 科学基础知识源和技术热点的确定

### 2.2 知识转移模型分析流程

知识转移模型分析分成 5 个步骤进行，如图 2-2 所示。

- 1) 专利检索，获取研究领域的目标专利数据集；
- 2) 相关数据读取，主要读取目标专利数据集的引用信息和施引信息；
- 3) 共被引聚类分析，包括论文和专利的共被引聚类分析，最后分别得到论文簇（科学基础知识源）和专利簇（技术热点）；
- 4) 知识转移的表征，包括知识转移路径和强度的建立；
- 5) 可视化，包括论文簇（科学基础知识源）和专利簇（技术热点）以及它们之间的关系。

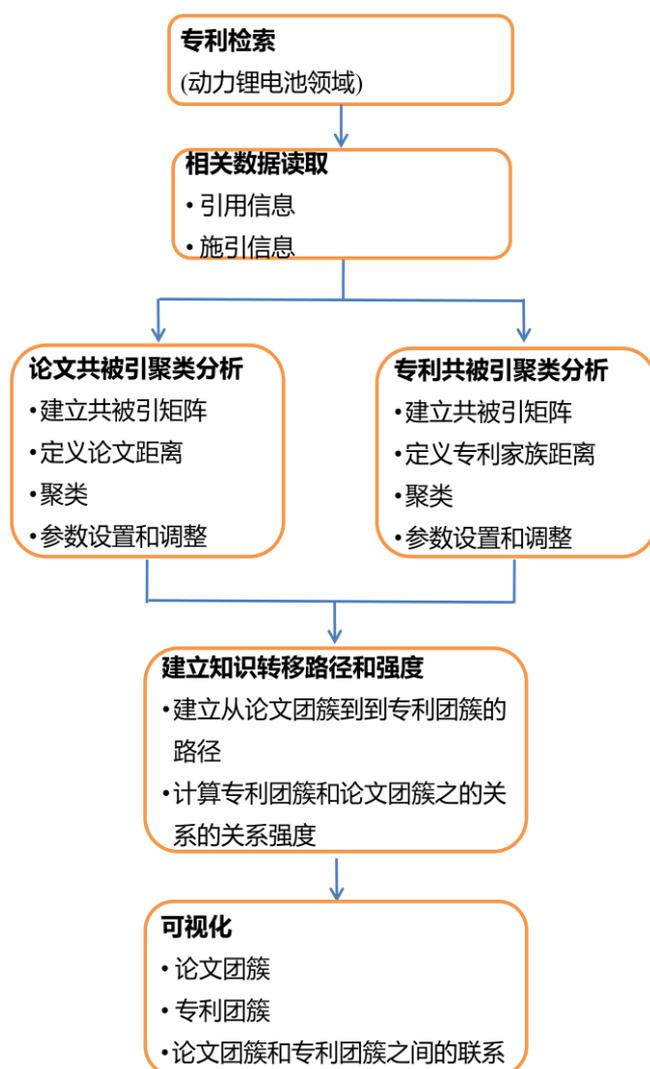


图 2-2 知识转移分析流程

### 2.3 数据获取

以 DII 数据库作为数据来源，但 DII 数据库只能提供引用的非专利文献的批

量下载，而不能批量下载施引专利。Thomson Innovation (TI) 中集合了 DII 专利数据，同时又能进行引用的非专利文献和施引专利的批量下载，如图 2-3 所示。所以本研究将在 DII 数据库中检索得到的专利数据的专利公开号 (PN) 导入 TI 中重新检索，获取 DII 同族信息，然后下载数据。



图 2-3 数据获取途径

专利引文由两部分组成。一部分是由申请人撰写专利说明书时，为了说明某项技术的发展历史及现状给出的参考引文，这部分引文嵌在说明书原文之中，不易摘取。另一部分是由审查员在对该发明进行三性审查时，检索出来的与本发明有关的关键文献，它们被清晰明确地著录在说明书的扉页上。

专利引文分析经常是以审查员给出的参考引文为基础。审查员参考引文与申请人参考引文具有很大的重复性。而且，审查员参考引文的数量远远大于申请人参考引文数量，以审查员参考引文作为专利引文分析的基础，不论从质量上还是从数量上都可以保证分析结果的可靠性。专利文件中所附的引文包括其它专利、科学论文、专业技术论文、书籍、文摘、学位论文等，通常，除专利外的引文统称为“非专利引文”。以非专利文献中对象，非专利文献包括科学杂志论文、会议论文、文摘、书籍，以及工业标准、工程手册等种种出版物。其中，科学杂志论文、会议论文、文摘更能反映基础研究的状况，本研究分析采用这些论文作为共被引聚类分析对象，得到基础知识源。

专利引用的论文在共被引聚类分析之前需要进行清洗和标准化，因为这部分

文献未经过加工。这些论文的详细信息需要通过 Web of Science 或者其他期刊进行查找和确认，最终完成对科学基础知识源的内容解读。

## 2.4 高被引论文或专利

如果目标数据集数据量过于庞大，那么选择高被引的论文或专利进行分析则至关重要。通常，被引用次数高的论文或专利被认为更具重要性。任意元素*i*，可以是论文或者专利。其被引用次数可以表征为 $\{c_{ik}\}$ ， $k = 1, \dots, n$ （施引元素）， $c_{ik}$ 取值为 1 或者 0。当*i*被*k*引用时， $c_{ik}$  为 1，反之为 0。元素的总被引次数为 $C_i = \sum_k c_{ik} = \sum_k c_{ik}^2$ 。所以，可以通过设置被引用次数阈值来选择高被引论文或者专利进行共被引聚类分析。另外，还可以根据引用分数法来选择论文或者专利进行共被引分析。元素*i*的引用分数 $F_i = \sum_k c_{ik} / T_k$ ，其中 $T_k = \sum_i c_{ik}$ 是元素*k*所有引用的文献数量<sup>15</sup>。

## 2.5 共被引关系的建立

以技术创造为出发点进行数据处理和分析，一项技术创造对应数据中的一个专利家族（一个专利家族中包含有若干件专利，专利之间互为同族专利），即研究对象的最小单元为一个专利家族（F）。TI 中获取的数据也是基于每个专利家族的引用的非专利文献（L）和施引专利（CP）信息。

对于引用的论文（L），如果专利 P1 和 P2 分别引用了论文 L1 和 L2，而专利 P1 和 P2 为同族专利，同属专利家族 F，那么 L1 和 L2 即为共被引关系，如图 2-4 所示。

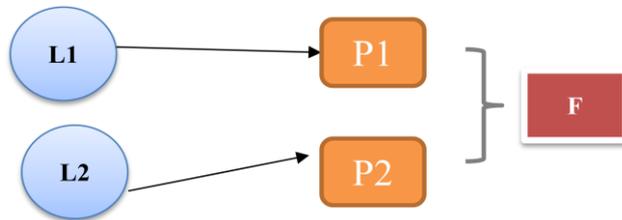


图 2-4 论文的共被引关系

对于专利家族(F),如果专利家族成员 F1 和 F2 分别被施引专利 CP I 和 CP II 引用，而 CP I 和 CP II 又互为同族专利，同属于另一个专家家族 Fc，则认为专

利 F1 和 F2 为共被引关系，如图 2-5 所示。施引专利之间同族关系的确定需将全部施引专利 CP 的专利号在 DII 中再进行检索，得到同族专利信息，建立同族专利对应关系。

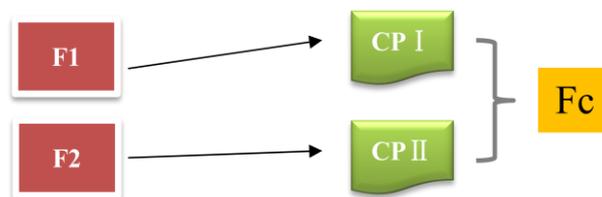


图 2-5 专利的共被引关系

## 2.6 共被引聚类分析

科学基础知识源和技术热点分别通过层次聚类法分析<sup>15-16</sup>得到。任意两个元素 $(i, j)$ ：可以是两篇被引论文，也可以是两个专利家族。它们的共被引次数为 $c_{ij} = \sum_k c_{ik} c_{jk}$ 。则二者的共被引强度根据 Salton 公式<sup>15</sup>可定义为：

$$\omega_{ij} = \frac{c_{ij}}{\sqrt{c_i + c_j}}$$

从而，两个元素之间距离定义为：

$$d_{ij} = 1 - \omega_{ij}$$

即： $i$ 和 $j$ 的共被引次数越大，则二者之间的距离越小。 $D_{ij}$ 最小值为 0，最大值为 1。

对于任意两个团簇 $\alpha$ 和 $\beta$ ，可以是论文团簇也可以是专利团簇，它们之间的距离类似两元素之间的距离。如果是单连接聚类，那么 $d_{\alpha\beta} = \min\{d_{ij}\}$ ，否则两个团簇的平均距离 $d_{\alpha\beta}$ 可以定义为：

$$d_{\alpha\beta} = \sum_{i \in \alpha, j \in \beta} d_{ij} / N_{\alpha\beta}$$

其中， $N_{\alpha\beta}$ 为两个团簇中的元素 $(i, j)$ 所有组合数目。

聚类过程为每次把距离最小的两个团簇进行合并。如果初始为  $N$  个元素，可以是被认为是  $N$  个团簇，每个团簇只包含 1 个元素。进行一步合并后成为 $(N-1)$ 个团簇，以此类推，最后形成稳定的聚类团簇。

$d_{\alpha\beta}$ 取不同的阈值将会得到不同的聚类结果。如果阈值取值过低，将会得到过多的聚类团簇，这些团簇过于密集不利于辨认和判读。如果阈值取值过高，将会得到很少的聚类团簇，这对聚类结果的届时必然存在局限性。因此，团簇之间

的距离 $d_{\alpha\beta}$ 进行反复的调试，最终选取合适的阈值。

共被引聚类分析最终将得到科学基础知识源（论文团簇）和技术热点（专利团簇）。合理的 $d_{\alpha\beta}$ 的设定，科学基础知识源和技术热点能够反映研究领域的技术创新和基础研究的主体结构。

## 2.7 知识转移表征

引用的非专利文献共被引聚类得到的簇即为科学基础知识源，专利共被引聚类得到的簇即为技术热点。建立科学基础知识源和技术热点的联系，即可发现知识是如何转移的。显然，路径和强度是知识转移过程中的两个重要参数。

根据技术热点和科学基础知识源之间的引用关系建立知识转移路径，如图 2-6 所示。如果 L1 被 F1 引用，则科学基础知识源与技术热点存在知识转移关系。

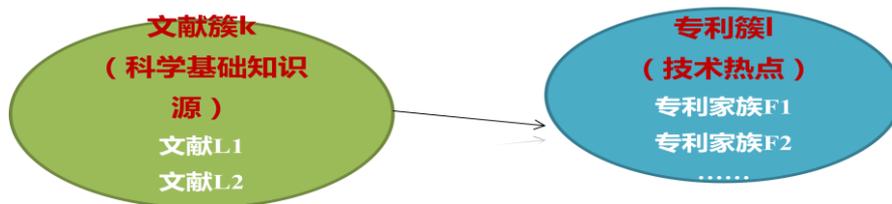


图 2-6 技术热点与科学基础知识源之间的关系

从科学基础知识源到技术热点的知识转移强度与技术热点的专利引用科学基础知识源中论文的次数、技术热点包含的专利数量、科学基础知识源包含的论文数量相关。知识转移强度定义为：

$$S_{AB} = \frac{K_{AB}}{\sqrt{M(A) \times N(B)}}$$

其中， $K_{AB} = \sum_{i \in \alpha, j \in \beta} c_{ij}$  是技术热点 A 中所有专利引用科学基础知识源 B 中所有论文的次数；M 是技术热点包含的所有专利的数量；N 是科学基础知识源包含的所有论文的数量。知识转移强度可以衡量技术创新和基础研究之间的关系强度。引用次数越多，关系越强，则科学基础知识源对技术热点的知识支撑强度越大，知识转移关系越强。

如果考量某个技术热点对科学基础知识的依赖程度或者某个科学基础知识源对技术创新的支撑程度，那么针对某个技术热点和科学基础知识源的总体知识

转移强度可以被定义为：

$$S_A = \frac{\sum_B K_{AB}}{\sum_B \sqrt{M(A) \times N(B)}}$$

## 2.8 可视化

通过重力模型<sup>17</sup>对论文团簇（科学基础知识源）和专利团簇（技术热点）分别进行可视化。重力模型假设每个元素同时受到引力和斥力的作用，当所受的力达到平衡时，元素之间的位置保持相对稳定。

如果选取科学基础知识源或者技术热点的数目为  $M$ ，在模拟的开始阶段，每个科学基础知识源或者技术热点分别被当成一个点，随机并且均匀地分布在  $(0,0)$  到  $(1,1)$  的空间区域内，由此可以计算任意两个点的空间距离  $r_{\alpha\beta}$ 。根据前面计算的团簇距离  $d_{\alpha\beta} (\alpha \in M, \beta \in M)$ ，来定义科学基础知识源或技术热点之间的相互作用力，其中，引力为：

$$F_{\alpha\beta}^- = d_{\alpha\beta} r_{\alpha\beta}$$

即，任意两个科学基础知识源或技术热点之间的引力由共被引强度和空间距离共同决定，两个科学基础知识源或技术热点之间共被引强度越高或者空间距离越远，则他们之间的引力越大，相互越靠近。

斥力为：

$$F_{\alpha\beta}^+ = d_{max} r_{\alpha\beta}^2$$

即，任意两个科学基础知识源或技术热点之间的斥力由最大共被引强度和他们之间的空间距离确定。空间距离越大，两个科学基础知识源或技术热点之间的斥力就越大。引力和斥力均沿两个点的连线方向。

科学基础知识源或者技术热点的理想平衡位置为

$$F_{\alpha\beta}^+ = F_{\alpha\beta}^-$$

$$\text{即： } r_{\alpha\beta} = d_{\alpha\beta} / d_{max}。$$

最后，科学基础知识源、技术热点、以及从科学基础知识源到技术热点的知识转移路径和强度利用 Raster3D<sup>∇</sup>分析工具进行展现，形成二维的知识转移地图。

<sup>∇</sup> Raster3D is a set of tools for generating high quality raster images of proteins or other molecules.

### 三、实证分析——转基因水稻

#### 3.1 转基因水稻简介

水稻是世界主要的粮食作物之一，世界上有将近二分之一的人口是以稻米为主食。然而在水稻的生产过程中由于各种病虫害、不良气候与环境的影响，严重制约了水稻的高产、稳产。随着转基因技术的兴起，转基因水稻的研究也成为了各国转基因技术研究的热点。1988年第一批转基因水稻<sup>18-19</sup>问世之后，越来越多的研究者利用转基因技术对传统的水稻进行改良，成功地获得了许多具有高产、抗性、营养乃至药用价值的转基因水稻<sup>20</sup>。转基因水稻根据其目的可分为三类：具有抗性的转基因水稻、高产和优质转基因水稻、具有药用价值的转基因水稻（详见表1）。

表 3-1 主要转基因水稻性状及对应基因\*

转基因性状		基因
一级	二级	
具有抗性	抗虫	苏云金杆菌杀虫结晶蛋白基因（BT 基因）：CRYIA 基因
		蛋白酶抑制剂（PI 基因）：豇豆胰蛋白酶抑制剂基因（CPTI 基因）、抗线虫的半胱氨酸蛋白酶抑制剂基因（OC-PIN2、SKIID86）、PIN2、SKTI
		抗稻飞虱等职务凝集素基因：雪花莲凝集素（GNA）基因、豌豆外源凝集素基因（P-LEC）
		其他如 SCIP-1、AHAS 等
	抗逆	抗涝基因：PDCI, PDCII, PDCIII
		抗旱基因：CMO 基因、SP17
		耐碱基因：CODA
		耐盐基因：BETA、BAD 的突变基因（*F*ADGH）H、乙烯受体类蛋白基因
		抗寒抗冻基因
	抗除草剂	抗草丁膦（phosphinothricin）的 BAR 基因
		抗草甘膦（glyphosate）的 EPSPS 基因

\*参考自：黄德林，2007，转基因水稻战略研究，中国农业出版社。

		原卟啉原氧化酶 (protox) 基因
		磷酸烯醇丙酮酸羧激酶 (PEPC) 基因、PCKTYPE 基因
		抗咪唑啉酮类除草剂以及其他
	抗病	水稻条纹叶枯病毒外壳蛋白基因 RSVCP
		抗白叶枯 Xa21 基因
		抗稻瘟病烟草几丁质酶基因
高产优质	品质性状改良	增加必须氨基酸 (如赖氨酸、甲硫氨酸): 水稻储藏蛋白基因 (GLUTILIN)
		增加维生素含量: 八氢番茄合成酶及其脱氢酶基因、PROLAMIN4A、PSY、CNTL、LCY 增加类胡萝卜素
		降低稻米的直链淀粉含量、AMYLOPULLULANASE
		增加总蛋白质含量: 大豆球蛋白基因
		维生素合成基因
	丰产	二磷酸核酮糖羧化酶基因 (RUBPC)、钙调素蛋白激酶基因以及 OSCEN1/2 基因
		ADPG 焦磷酸化酶的突变基因 (GLGC-TM)、AGP 淀粉合成酶基因
		磷酸蔗糖合成酶基因
		分蘖等其他基因
	其他	药用蛋白、抗体、疫苗等基因
与生理生长 (花色、激素以及矮化等) 有关基因		
不育相关基因、花粉特异表达不育基因		

### 3.2 转基因水稻技术领域的科学基础知识源

截止 2012 年 6 月 12 日,从 TI 中基于 DWPI 专利数据库共获得 4274 项专利,即 4274 条专利家族记录。共有 19,372 篇论文被转基因水稻领域的专利所引用,对所有被引用的论文进行共被引聚类分析。设定论文团簇距离 $d_{ij}$ 阈值为 0.90,被引论文共被引聚类最后形成 10 个科学基础知识源,每个科学基础知识源包含的论文数量超过 40 篇,最大的一个科学基础知识源包含的论文数量为 792 篇。通过专家咨询,对每个科学基础知识源的研究内容作出解读。表 3-2 给出了转基因水稻技术创新所借鉴的 10 个科学基础知识源的详细信息。

表 3-2 转基因水稻领域科学基础知识源相关信息

编号	研究内容	论文数量	总体知识转移强度
S1	Research on important enzymes in transgenic plants, which includes protoporphyrinogen oxidase as target for herbicide, APS reductase, acetyl coenzyme A carboxylase, etc.	65	0
	转基因植物涉及的重要或关键酶研究, 包括作为除草剂靶标的原卟啉原氧化酶、APS 还原酶、乙酰辅酶 A 羧化酶等等		
S2	Research on genetic characterization.	62	0.001
	基因特征相关研究		
S3	Research on important physiological processes.	87	0.001
	重要生理过程研究		
S4	Research methods of gene function.	53	0.025
	基因功能研究方法		
S5	Research on genetic stability of exogenous gene in plants.	97	0.007
	外源基因在植物中的遗传稳定性研究		
S6	Cloning, sequence characteristics, comparison and evolution of the important agronomic trait genes and the insecticidal proteins encoding genes, etc.	573	0.342
	重要农艺性状基因以及杀虫蛋白编码等基因等克隆、序列特征、比较及进化		
S7	Research on insect-resistant, diseases-resistant, stress-resistant transgenic plants.	44	0.007
	抗虫、抗病、抗逆转基因植物研究		
S8	Research on genetic transformation in plants.	259	0.270
	植物遗传转化体系研究		
S9	Research on gene regulation in plants.	792	0.329
	植物基因调控研究		
S10	Research on gene transcription in plants.	41	0.017
	植物基因转录研究		

从转基因水稻领域科学基础研究对技术创新的支撑角度看, S6 (重要农艺性

状基因以及杀虫蛋白编码等基因等克隆、序列特征、比较及进化)、S9 (植物基因调控研究) 和 S8 (植物遗传转化体系研究) 等三个科学基础知识源的总体知识转移强度相对较大, 分别为 0.342、0.329 和 0.270。可见, 这三个基础知识源对转基因水稻的技术创新的支撑强度最大。S1 (转基因植物涉及的重要或关键酶研究)、S2 (基因特征相关研究) 和 S3 (重要生理过程研究) 虽然是转基因水稻领域的科学基础知识源, 但是它们对技术热点却没有明显的知识转移趋势。这可能是因为这三个科学基础知识源所支撑的技术创新并不是转基因水稻领域的技术热点。造成此结果的原因还有可能是这三个科学基础知识源所支撑的技术创新是转基因水稻领域新出现的研究方向, 而这个研究方向正处于开始阶段, 尚未成为转基因水稻领域的技术热点, 这需向领域专家进行更深一步的咨询。

### 3.3 转基因水稻技术领域的技术热点

根据检索得到的转基因水稻领域的相关专利被后续专利所引用的情况, 对转基因水稻领域的 4274 项专利进行共被引聚类分析。设定专利团簇距离 $d_{ij}$ 阈值为 0.86, 专利共被引聚类最后形成 8 个技术热点, 每个技术热点包含的论文数量超过 20 项, 最大的一个技术热点包含的专利数量为 387 项。通过专家咨询, 对每个技术热点的技术内容作出解读。表 3-3 给出了转基因水稻技术创新所形成的 8 个技术热点的详细信息。

表 3-3 转基因水稻领域技术热点相关信息

编号	技术内容	专利数量	总体知识转移强度
T1	Development in new or modified nucleotide, which can help transgenic rice to obtain new characters or improve genetic transformation efficiency.	35	0.005
	用于转化水稻的新的或者修饰核苷酸片段, 以使转基因植物获得新性状或者提高转化效率等		
T2	Development in new nucleotide resistant to herbicides such as glyphosate, etc.	29	0
	具有抗草甘膦等除草剂性状的新核苷酸片段		
T3	New promoter sequences for new characters or increasing	37	0.016

	yield.		
	新启动子序列, 用于转化植物以产生新形态或增产等		
T4	New nucleotide or recombinant DNA vector resistant to biological and abiotic stress.	36	0.008
	新核苷酸分子、新重组 DNA 载体, 以使转基因植物具有抗生物和非生物胁迫的特性		
T5	New regulatory factors and method of changing nucleotide sequence, whose purpose is to improve insect-resistance, diseases-resistance, high plant density, nutritional value or content of amino acids, etc. in transgenic rice.	387	0.722
	新调控因子、或通过改变核苷酸序列等生产转基因水稻的方法, 其目的是或提高抗虫性、抗病性、植物在高密植的生长、或提高营养价值		
T6	Separating and obtaining the origins of gene, which includes new strains for pest and disease control, new plant cells resistant to herbicides, etc.	50	0.017
	分离或获取基因来源, 包括用于防治植物害虫、病害等的新菌株、耐受除草剂新植物细胞等等		
T7	Production methods of transgenic rice with insect-resistance, stress-resistance or high yield.	48	0.107
	生产抗逆、抗虫或高产的转基因水稻新方法		
T8	New compounds or compositions for weed, pest and disease control, which should be capable to kill weeds, pests or diseases, etc., but not to be harmful to the growth of transgenic rice.	117	0.125
	用于防治杂草、病害和虫害等的新化合物或组合物		

从转基因水稻领域技术创新对科学基础研究的依赖角度看, T5(新调控因子、或通过改变核苷酸序列等生产转基因水稻的方法)、T8(用于防治杂草、病害和虫害等的新化合物或组合物)和 T7(生产抗逆、抗虫或高产的转基因水稻新方法)等三个技术热点的知识转移活动相对频繁, 其总体知识转移强度分别为 0.722、0.125 和 0.107。T2(具有抗草甘膦等除草剂性状的新核苷酸片段)与科学基础研究的联系较弱, 技术创新过程中没有明显的知识转移活动。

### 3.4 转基因水稻技术领域的知识转移

图 3-1 为转基因水稻技术领域的知识转移地图。其中，绿色圆代表技术热点，黄色圆代表科学基础知识源。圆的大小正比于技术热点或科学基础知识源中所包含的专利家族或论文数量；同颜色圆之间的相对位置反映它们之间的关联程度，距离越近，关联程度越高。技术热点和科学基础知识源之间的连线表示存在知识转移活动，连线的粗细正比于知识转移强度。图中只展现知识转移强度高于 0.03 的知识转移。

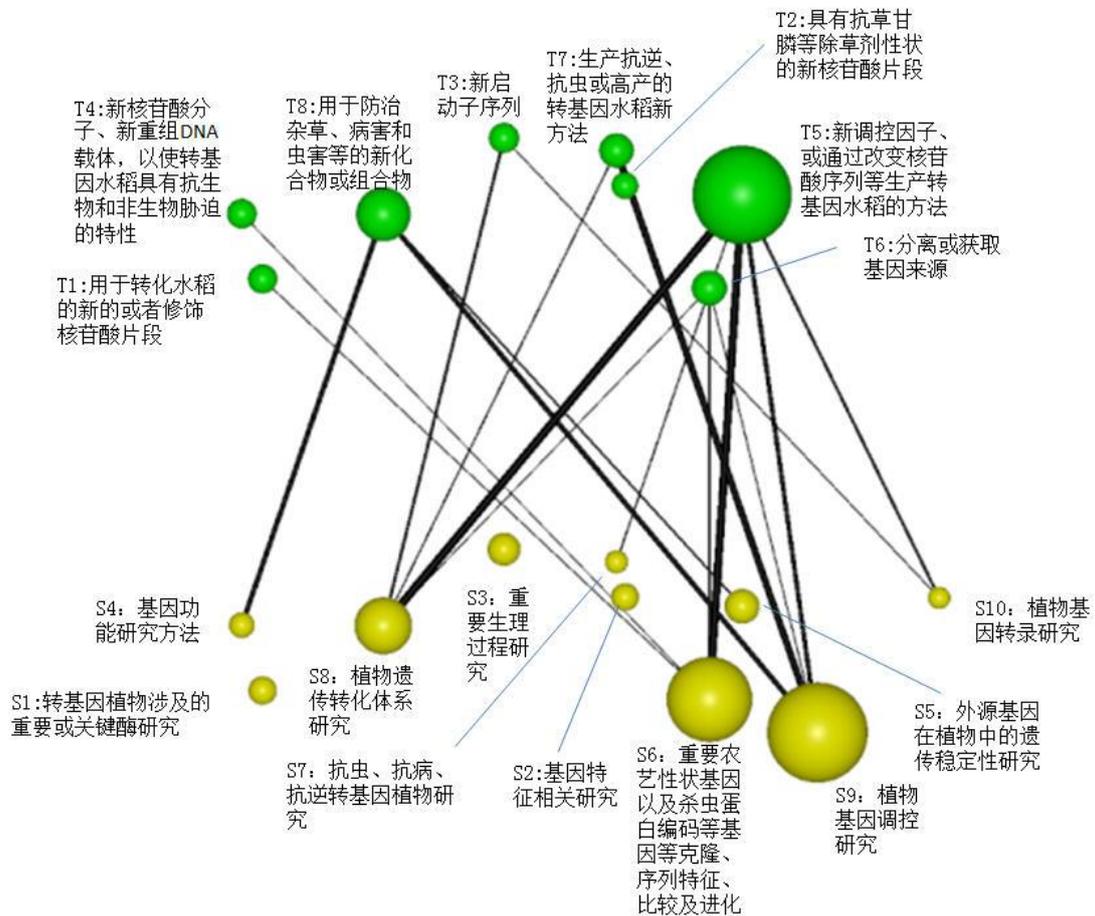


图 3-1 转基因水稻技术领域的知识转移地图

转基因水稻技术领域中，技术热点与科学基础知识源存在着不同程度的联系。技术热点 T5（新调控因子）主要依赖于多个科学基础知识源：S8（植物遗传转

化体系研究), 知识转移强度为 1.22; S6 (重要农艺性状基因以及杀虫蛋白编码等基因等克隆、序列特征、比较及进化), 知识转移强度为 1.06; S9 (植物基因调控研究), 知识转移强度为 0.35; S10 (植物基因转录研究), 知识转移强度为 0.17; S7 (抗虫、抗病、抗逆转基因植物研究), 知识转移强度为 0.07。T8 (用于防治杂草、病害和虫害等的新化合物或组合物) 的技术创新主要来源于基础知识研究 S9(植物基因调控研究), 知识转移强度为 0.47; S4(基因功能研究方法), 知识转移强度为 0.46; S5 (外源基因在植物中的遗传稳定性研究), 知识转移强度为 0.10。技术创新 T7 (生产抗逆、抗虫或高产的转基因水稻新方法) 主要借鉴了科学基础研究 S9 (植物基因调控研究), 知识转移强度为 0.79; S (植物遗传转化体系研究), 知识转移强度为 0.07。技术创新 T6 (分离或获取基因来源) 的借鉴的科学基础研究是 S6 (重要农艺性状基因以及杀虫蛋白编码等基因等克隆、序列特征、比较及进化), 知识转移强度为 0.06; S9 (植物基因调控研究), 知识转移强度为 0.05; S8 (植物遗传转化体系研究), 知识转移强度为 0.04。技术创新 T3 (新启动子序列) 所依赖的科学基础知识源是 S8 (植物遗传转化体系研究), 知识转移强度为 0.17; S10 (植物基因转录研究), 知识转移强度为 0.05。技术创新 T1 (用于转化水稻的新的或者修饰核苷酸片段) 和 T4 (新核苷酸分子、新重组 DNA 载体, 以使转基因水稻具有抗生物和非生物胁迫的特性) 均借鉴了科学基础知识研究 S6 (重要农艺性状基因以及杀虫蛋白编码等基因等克隆、序列特征、比较及进化), 知识转移强度分别为 0.05 和 0.04。

## 四、结论

### 4.1 研究成果

从基础研究到技术创新的知识转移过程研究一直以来颇受关注。专利引文分析可以利用专利来分析基础研究和技术创新的联系, 为深入探索基础研究对技术创新的影响提供了一种方法。本项目提出了基于专利引文的知识转移分析模型, 探索技术领域的知识转移情况, 了解技术创新与科学基础研究的联系。该模型避免了传统知识转移研究方法缺乏普适应、指标难以衡量等局限性, 适用于所有技

术领域的知识转移研究。主要研究成果包括：

#### 1、建立知识转移分析模型

基于专利引文，利用共被引聚类分析方法，本项目建立了知识转移分析模型。通过知识转移分析模型，可以揭示研究领域内的技术热点、科学基础知识源，并且表征了二者之间的知识转移路径和强度，最终形成知识转移地图。该分析模型具有较强的实用性和可操作性，适用于所有技术领域的知识转移研究。知识转移分析模型立足于专利文献公开信息，侧重于定量化手段揭示领域内的知识转移过程，在专利分析中具有一定的参考价值。

#### 2、揭示了转基因水稻技术领域的知识转移

以转基因水稻技术领域为研究对象进行知识转移模型的实证分析，得到了较为合理且客观的该领域知识转移的分析结果。从科学基础研究 S8（植物遗传转化体系研究）和 S6（重要农艺性状基因以及杀虫蛋白编码等基因等克隆、序列特征、比较及进化）到技术热点 T5（新调控因子）、从科学基础研究 S9（植物基因调控研究）到技术创新 T7（生产抗逆、抗虫或高产的转基因水稻新方法）的知识转移是转基因技术领域内最为活跃的知识转移。

#### 3、形成知识转移分析工具

本项目形成了一套知识转移分析工具，在该分析工具中对分析的指标和参数的设置均进行了详细说明，可为今后同类工作借鉴。该分析工具将在情报研究与服务实践中不断优化和完善。

### 4.2 存在问题

由于受到专利引文信息的限制，为研究工作带来一些不可避免的局限之处。目前，尚无数据库对专利所引用的非专利文献进行深加工，因而这些数据相对杂乱，虽然项目研究中探索的规则可以完成一部分清洗工作，但需手工清洗的数据还是需要大量的工作量。

### 4.3 未来研究方向

在现有项目研究成果基础之上，还可就下述方向开展进一步的深入探索：

1) 继续选取若干领域（例如信息领域、材料领域等），分别进行实证分析，比较研究这些领域知识转移过程。并且，在实证分析中，进一步完善模型，确保模型在大数据量条件下可运行，并且验证分析结果的合理性。

2) 探索技术热点对主要科学基础知识源的依赖程度与专利重要性的关系，并且总结规律。

3) 从引文的期刊信息出发，探索学科领域对技术创新的支撑。

4) 从国家信息入手出发，探索从基础知识到技术创新的知识转移国家网络。

5) 从机构信息出发，探索从基础知识到技术创新的知识转移国家网络。

6) 从机构类型入手出发，探索从基础知识到技术创新的知识转移部门网络。

## 参考文献:

- [1] 赵黎明、高杨、韩宇 (2002), 专利引文分析在知识转移机制研究中的应用, 科学学研究, 20 (3), 297-300.
- [2] 林岩 (2008), 基于专利数据的知识计量研究评述, 科技管理研究, 9, 91-93.
- [3] 韩宇, 知识转移及其分析方法研究, 天津大学博士学位论文
- [4] Jaffe, A. B., Fogarty, M. S., Banks, B. A. (1998), Evidences from patents and patent citations on impact of NASA and other federal labs on commercial innovation, *Journal of Industrial Economics*, 183-205.
- [5] Narin F., et al. (1997), The increasing linkage between U.S. technology and public science, *Research Policy*, (26), 317-330.
- [6] Narin, F., Mbert, M., Kroli, p., Hiexs, D. (2000), Inventing our future—link between Australian patenting and basic science, CHI Research Inc.
- [7] Meyer, M. (2002), Tracing knowledge flows in innovation systems. *Scientometrics*, 54(2), 93-212.
- [8] Chen C, Hicks D. (2004), Tracing knowledge diffusion, *Scientometrics*, 59(2), 199-211.
- [9] [http://foresight.jst.go.jp/en/jst\\_indicators/tech\\_linkage/](http://foresight.jst.go.jp/en/jst_indicators/tech_linkage/)
- [10] The section of "Connecting to Knowledge" of the "OECD Science, Technology and Industry Scoreboard 2011".
- [11] Small, H. (1973). Co-citation in Scientific Literature: A New Measure of Relationship Between Two Papers. *Journal of the American Society for Information Science*, 24 (4), 265-269.
- [12] Moge, M. E. (1991). Patent co-citation analysis of Eli Lilly & Co. patents. *Research-Technology Management*, 34(4), 43-49.
- [13] Lai, K. K., & Wu, S. J. (2005). Using the patent co-citation approach to establish a new patent classification system. *Information Processing & Management*, 41(2), 313-330.
- [14] Wang, X.W., zhang, X., & Xu, S.M. (2011) Patent co-citation networks of Fortune 500 companies. *Scientometrics*, 88, 761-770.

- [15] Small, H. & Sweeney, E. (1985a). Clustering the Science Citation Index Using Co-citations. I . A Comparison of Methods. *Scientometrics*, 7(3-6), 391-409.
- [16] Small, H., Sweeney, E. & Greenlee, E. (1985b). Clustering the Science Citation Index Using Co-citations. II . Mapping Science. 8(5-6), 321-340
- [17] Neil R. Quinn JR & Melvin A. Breuer. (1979). A Forced Directed Component Placement Procedure for Printed Circuit Boards. *IEEE Transactions on Circuits and Systems*, 26(6), 377-388.
- [18] Toriyama, K., Arimoto, Y., et al. (1988). Transgenic Rice Plants after Direct Gene Transfer Intoprotoplasts. *Bio/Technology*. 6, 1072-1074.
- [19] Zhang, W., Wu., R. (1988). Effieient Regeneration of Transgenic Plant from Rice Protoplasts and Correctly Regulated Expression of the Foreign Gene in the Plant. *Theoretical and Applied Genetics*. 76(6), 835-840.
- [20] 胡贻椿、陈天金、朴建华、杨晓光（2009），转基因水稻及安全性的研究进展，*中国食物与营养*，8，19-22.