

# International Research Status Analysis of Tomato Molecular Breeding Based on Literature Survey

Sufen Sun<sup>1,2</sup>, Changshou Luo<sup>1,2,\*</sup>, Qingfeng Wei<sup>1,2</sup>, Junfeng Zhang<sup>1,2</sup>

<sup>1</sup>Institute of Information on Science and Technology of Agriculture, Beijing Academy of Agriculture and Forestry Sciences, Beijing, China

<sup>2</sup>Rural Distance Information Service Engineering Technology Research Center of Beijing Beijing, China

## Email address:

[sunsf@agri.ac.cn](mailto:sunsf@agri.ac.cn) (Sufen Sun), [luochangshou@163.com](mailto:luochangshou@163.com) (Changshou Luo), [weiqingfeng201@163.com](mailto:weiqingfeng201@163.com) (Qingfeng Wei),

[zhangjf@agri.ac.cn](mailto:zhangjf@agri.ac.cn) (Junfeng Zhang)

\*Corresponding author

## To cite this article:

Sufen Sun, Changshou Luo, Qingfeng Wei, Junfeng Zhang. International Research Status Analysis of Tomato Molecular Breeding Based on Literature Survey. *American Journal of BioScience*. Vol. 4, No. 4, 2016, pp. 53-57. doi: 10.11648/j.ajbio.20160404.13

Received: June 11, 2016; Accepted: August 26, 2016; Published: September 2, 2016

**Abstract:** The international research trend of tomato molecular breeding was studied based on thesis analysis in SCIE. Through the analysis of the annual number of articles, reveals the variation trend of related research. By the analysis of Scientific productivity and achievement influence, compares the research competitiveness among countries. Use the co-word clustering to identify 6 important research topics in the field, and use the strategy coordinates to evaluation the relative position of these topics .The result shows that, tomato molecular breeding research presents a rising trend. The United States plays an absolute leading role in the field. China's research layouts more but achievements influence is relatively low. # 5 (tomato genome) and # 1 (resistance) themes are the center of tomatoes molecular breeding. # 2 (oxidative stress) may be the center of the future research.

**Keywords:** Tomato, Molecular Breeding, Situation, Co-word Analysis, Strategic Coordinates

## 1. Introduction

Tomatoes originated in South America. It has the long history of cultivation, can adapt to a wider scope, and is the most widely grown vegetable in China and worldwide. It can be eaten, and as a medicinal and ornamental plants. It is an important cash crop in China. In the basic scientific research, it has advantages of small genome, short reproductive cycle, and easy cultivation and transgenesis, and the kinship is fare with Arabidopsis, rice, corn etc., gradually become a model system of fruit developmental mechanism. Its findings of genetic theory, is often used to draw on other crops [1]. Because tomatoes susceptible to disease, also climacteric fruit, intolerance storage, perishable. Therefore, breeding disease resistance and shelf-stable varieties, is the focus of each breeding companies and units [2]. In addition, cold resistant, drought tolerant, salt tolerance, heat, detoxification is an important content of tomato resistance breeding; Sweetness increase, the contents of soluble solids and lycopene is an important content of tomato quality breeding.

According to the need of yield and quality of tomato, molecular marker, gene engineering and molecular breeding technology provides a good solution. The application of these techniques can reduce the blindness of breeding, the breeding goals more accurately, shorten the breeding period, which greatly increases the speed of new variety breeding. The current related research mainly concentrated in disease-resistant [3] [4] [5], resilience [6] [7], quality improvement [8], delay fruit ripening and preservation [9], to create male sterile [10], herbicide resistant [11] and so on several aspects.

## 2. Data Source and Search Strategy

Take SCIE (Science Citation Index Expanded) as data source. The coverage of SCIE is wide. It focuses on the basic research of various disciplines and includes the high impact journal. The database paper quality is high.

By analysis the key technology of tomato molecular breeding, and consult the field expert to construct the search term:

TI=(tomato OR "Lycopersicum esculentum" OR "L. esculentum" OR "Solanum lycopersicum" OR "S. lycopersicum") AND TS=((tomato OR "Lycopersicum esculentum" OR "L. esculentum" OR "Solanum lycopersicum" OR "S. lycopersicum") near/25[7] (breed OR cultivar OR strain OR line OR species OR variety OR plant)) AND TS=(marker OR MAS OR SSR OR AFLP OR CAPS OR RFLP OR RAPD OR InDel OR QTL OR SNP OR Transgenic OR Transgene OR gene OR genetic OR genome OR genotype OR phenotype OR GMO OR DNA OR RNA OR mRNA).

Use it above to obtain the result. Retrieve time until September 14, 2015.

### 3. The Research Status

#### 3.1. The Annual Trends

The number change of papers published in tomato molecular breeding field each year can reflect the development trend. The data change is shown in figure 2 [8]. From it, there are 3 stages of tomato molecular breeding research. 1965-1990 is the early stage, fewer papers. 1990 is the most, 9 papers. And annual change is very small. Growth rate is not obvious. 1991-2006 is the steadily growing stage. The number of publications every year have a jump compared with the previous stage, but in this period, it is

stable development relative. About 50-100 papers per year. Among them, 160 patents in 2005, it is the most in this phase. 2007-2012 is rapid growth stage. The paper number breaks through 200 in 2009 and reach the peak in 2012 (280 papers). Since then, it is falling stage, but still remain in the higher level of 250 papers or so. Mainly affected by 2012 tomato genome sequenced.

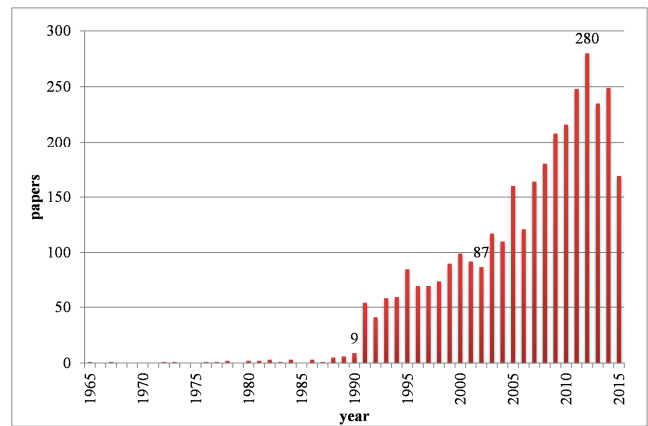


Figure 1. The annual distribution of paper number.

#### 3.2. Key Nation and Region Studies

By the retrieving data, tomato molecular breeding thesis involved 105 countries and regions. Table 2 is the top 15. It reflects the papers number, the first author papers number and cited frequency.

Table 1. The top 15 national publications and citation in tomato molecular breeding.

NO.	Country	Total papers			First author papers			The percentage of First author papers
		Number	Total cited	Average cited	Number	Total cited	Average cited	
1	USA	881	29021	32.94	745	25559	34.31	84.56%
2	China	288	3297	11.45	248	2049	8.26	86.11%
3	Spain	272	5908	21.72	224	4036	18.02	82.35%
4	Netherlands	229	8119	35.45	195	6626	33.98	85.15%
5	Italy	226	3171	14.03	187	1975	10.56	82.74%
6	France	213	6351	29.82	136	3971	29.20	63.85%
7	Japan	188	3168	16.85	154	2100	13.64	81.91%
8	Germany	184	6139	33.36	110	3519	31.99	59.78%
9	Israel	170	7940	46.71	122	5293	43.39	71.76%
10	Brazil	163	1603	9.83	149	966	6.48	91.41%
11	India	160	1120	7.00	137	811	5.92	85.63%
12	UK	158	6398	40.49	98	4425	45.15	62.03%
13	South Korea	69	323	4.68	61	201	3.30	88.41%
14	Canada	64	1914	29.91	49	1197	24.43	76.56%
15	Australia	59	1353	22.93	42	859	20.45	71.19%

America: The number of publications and the first author papers are the most, and far more than China which is ranked second. It indicates that America has played the absolute leading role in this field. The cited frequency of first author paper is 34.31. It is higher than the average citation rates of all papers (32.94). It reflects that the United States research results have a higher influence than other countries.

China: The number of publications and the first author

papers are ranked second. The cited frequency of all paper and first author paper is 11.45 and 8.26 respectively, and are ranked the twelfth. It reflects that the China research results have a lower influence.

Research of the Dominant: The proportion of first author papers reflects the nation's research dominance. Brazil, South Korea have the highest percentage which is 91.41% and 88.41% respectively. It shows that the two countries in this

field has a dominant position. Germany and UK have the lowest percentage which are 59.78% and 62.03%. It indicates that the two countries fewer take part in other national studies.

Paper influence. The average cited frequency of each paper can show its influence in the study. Israel has published 170 papers, the average cited frequency of each paper is 46.71, the influence is the highest. Among them, the first author paper is 122, the average cited frequency of each paper is 43.39, ranking the second. United Kingdom has published 158 papers, the average cited frequency of each paper is 40.49, the influence ranked second. Among them, the first author paper is 98, the average cited frequency of each paper is 45.15 times, ranking the second.

## 4. Subject Distribution and Characteristics

### 4.1. Theme Identification

Select the keywords from the title and abstract use the method of natural language procession, and form a set with the author keywords and machine keywords from the Web of Science. And select the phrase which frequency is greater than 10 times to co-occurrence analysis. It is a total of 217, involving paper 3049, accounting for 90.29% of all paper. Divided the phrase set into 6 topics. Screening outstanding phrases as the core theme of the topic. The results are shown in Table 2.

**Table 2.** The 6 topic of tomato molecular breeding.

Topic code	Keywords	Related papers	Percentage	H index
#1	disease resistance; pseudomonas syringae; defense responses; signaling pathways; cell death	958	28.37%	79
#2	oxidative stress; stress tolerance; abiotic stress; superoxide dismutase; lipid peroxidation	517	15.31%	49
#3	fruit quality; introgression lines; fruit development; fruit size; soluble solids content	1027	30.41%	73
#4	genetic diversity; fruit yield; Agrobacterium tumefaciens; simple sequence repeat; Alternaria solani	748	22.15%	53
#5	tomato genome; gene family; expression pattern; genome-wide analysis; plant development	568	16.82%	66
#6	linkage map; genetic mapping; short arm; nematode resistance gene; genomic region	778	23.04%	71

From the Table 2 above:

The largest theme #3: The theme is related to the tomato fruit of fruit quality, fruit development, fruit size, soluble solids content, etc. It has the most publications, a total of 1027 articles, it is 30.41% of all papers in this field.

The most influential theme #1: The theme is related to tomato disease resistance, such as the disease resistance, pseudomonas syringae, defense responses, signaling pathways, cell death, etc. The H index is 79, and ranked first of 8 topics in this field. In addition, it has 958 articles, it is 28.37% of all papers in this field, ranked the second.

### 4.2. Strategic Coordinates

Use the centrality and density to analysis the 6 research topics characteristics of tomatoes molecular breeding, so as to reflect the status and development prospects of every theme. Density is used to measure the connection strength of theme inside, it represents the theme cohesion and maturity. The denser the theme, the stronger the correlation, the more stable of the theme structure, the research is more mature. The centrality is used to measure the interrelated degree between theme, it represent the position of subject in the field. The greater the centripetal degree, the closer with other topics, the more central place in the whole field. The strategic coordinates of 6 themes of tomato molecular breeding are shown as table1:

**Table 3.** The strategic coordinates of 6 topics.

Topic code	density	centrality
#1	0.3945	0.1134
#2	0.3242	0.0897

Topic code	density	centrality
#3	0.2143	0.0727
#4	0.0792	0.0765
#5	0.3400	0.1395
#6	0.2257	0.1011

Take centrality as the abscissa, density as the ordinate, take the average of centrality and density as the origin, and draw strategy coordinate diagram. The result is shown in figure 2, all topics in four quadrants, and has different meanings:

The first quadrant, density and centrality are higher. It shows that the correlation within themes is stronger. It is a central theme in tomato molecular breeding, and get the favor of researchers for a long time. They are #1 (disease resistance; pseudomonas syringae; defense responses; signaling pathways; cell death) and #5 (tomato genome; gene family; expression pattern; genome-wide analysis; plant development).

The second quadrant, density is high, but centrality is low. It indicates that the internal correlation between the me is strong. It is the mature theme and non-core. They are #6 (linkage map; genetic mapping; short arm; nematode resistance gene; genomic region);

The third quadrant, density and centrality are low. It shows that the correlation within themes is loose, research system has not been formed, and the difference between research topic is great, research value needs further study. They are #3 (fruit quality; introgression lines; fruit development; fruit size; soluble solids content) and #4 (genetic diversity; fruit yield; Agrobacterium tumefaciens; simple sequence repeat; Alternaria solani).

The fourth quadrant, density is low, but centrality is higher.

It shows that the correlation within is loose, but closely linked with other themes. Its strategic position cannot be ignored. It is emerging, will soon become the subject of future research center, and with room for further development. They are #2 (oxidative stress; stress tolerance; abiotic stress; superoxide dismutase; lipid peroxidation).

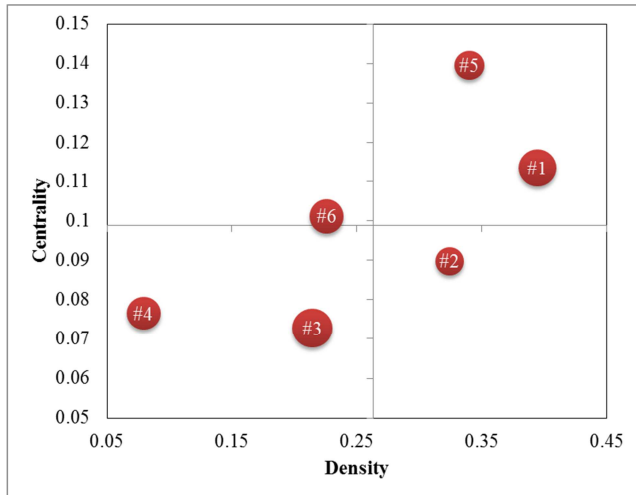


Figure 2. The annual distribution of paper number.

## 5. Conclusion

Tomato molecular breeding research shows a rising trend. According to the annual distribution of paper amount, it can be roughly divided into three stages: the early stage: from 1965 to 1990; the steady growth stage: from 1991 to 2006; the rapid growth stage: after 2007. Among them, the maximum amount to the relevant papers before and after 2012, with 2012 tomato genome sequenced. Among them, the paper publication reaches the peak in 2012 (280 papers). Since then, it is falling stage. It mainly affected by 2012 tomato genome sequenced.

US has a dominant position in tomato molecule breeding, and has a higher influence compared with other countries. China in this field layout more, but the results influence is low compared with other countries. The first author paper of Brazil and South Korea accounts for the highest percentage of their own papers respectively, research leading is strongest. The first author paper of Germany and the UK accounts for the lowest percentage of their own papers respectively, they are mostly involved in tomato molecular breeding with other countries. The paper cited frequency of Israel is the highest in the TOP15 countries. Their influence is the highest. The first author paper cited frequency of UK is the highest, and has the highest influence in the TOP 15 countries.

From the co-word clustering analysis, tomato molecular breeding research can be divided into 6 important topics. Among them, the largest theme is tomato fruit-related research (#3), such as the fruit quality, fruit development, fruit size, soluble solids content. The most influential theme is tomato disease resistance research (#1), such as the *Pseudomonas syringae*, defense responses, signaling pathways, cell death, etc.

Among the 6 topics, the inner correlation of #5 and #1 is strong, and at the center of the field. The inner correlation of #2 is weak, but closely linked with other themes, its strategic position shouldn't be regarded, it is likely to be the research center in future and worth focus on.

## Acknowledgements

The research work was supported by the National Sci-Tech Support Plan "Construction and application of provincial rural information service platform in developed area" (NO.2014BAD10B02), and the young scientist fund of BAAFS: Research on agricultural scientific technical online counseling system based on knowledge map (NO. QNJJ201534). Correspondence to Chang-Shou Luo (luocs@agri.ac.cn).

## References

- [1] Lin, T; Zhu, GT; Zhang, JH; et al. Genomic analyses provide insights into the history of tomato breeding [J]. NATURE GENETICS, 2014, 46 (11): 1220-1226.
- [2] Kong Fanhui. The study of creating multi-resistance and long shelf life tomato materials by marker-assisted selection [D]. Northeast Agricultural University, 2015.
- [3] ZHU Ming-tao. Research Progress on Molecular Breeding for Disease Resistance of Tomato [J]. Northern Horticulture, 2010, 23: 200-203.
- [4] Sim, Sung-Chur; Robbins, Matthew D.; Wijeratne, Saranga; et al. Association Analysis for Bacterial Spot Resistance in a Directionally Selected Complex Breeding Population of Tomato [J]. PHYTOPATHOLOGY, 2015, 105 (11): 1437-1445.
- [5] Santos, Jean C. S.; Siqueira, Walter J.; Melo, Paulo C. T.; et al. Selection of tomato breeding lines with resistance to Tomato yellow vein streak virus [J]. HORTICULTURA BRASILEIRA, 2015, 33 (3): 345-351.
- [6] Cao Yongxiang, Zhang Xichun. Molecule Marker Technology in The tomato Resistance Breeding Application [J]. Chinese Agricultural Science Bulletin, 2008, 02: 81-88.
- [7] Li Junming, Song Yan, Zhu Tong, et al. Reviews about Molecular Breeding of Salt Tolerance on Tomato [J]. Molecular Plant Breeding, 2006, 01: 111-116.
- [8] LI Xiao-lei, LI Jing-fu, KANG Li-gong, et al. Research Progress in Tomato Quality Genetics and Breeding [J]. CHINA VEGETABLES, 2010, (14): 1-3.
- [9] Bai Yan-jing. Selective breeding of transgenic anti-decayed tomato inbred lines and hybrid generation variety test [D]. Huazhong Agricultural University, 2013.
- [10] ZHANG He, LI Jing-fu, ZHUANG Lei, et al. Initial Exploration about Creation of New Tomato Male Sterile Line and Its Seed Production Technology [J]. China Vegetables, 2012, 06: 93-95+108.
- [11] WANG Ao-Xue, CHEN Xiu-Ling. Current status and industrialization of transgenic tomatoes [J]. Hereditas, 2011, 33 (9): 962-974.

- [12] Kostoff RN, Eberhart HJ, Toothman DR. Database tomography for information retrieval. JOURNAL OF INFORMATION SCIENCE, 1997, 23 (4): 301-311.
- [13] Yuan Jianxia, Dong Yu, Zhang Bo, et al. A Bibliometrical Analysis of International Development of Molecular Marker Assisted Breeding of Crop [J]. Science Focus, 2012, 02: 24-32.
- [14] ZHOU Li-ying; ZUO Wen-ge. Soil science research profiling based on subject co-occurrence [J]. Journal of China Agricultural University, 2014, 04: 265-272.