Journal of Informetrics 15 (2021) 101207

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Scientific collaboration and career stages: An ego-centric perspective

Wei Lu[a,b,1], Yan Ren[a,b,1], Yong Huang[a,b,*], Yi Bu[c], Yuehan Zhang[a,b]

[a] School of Information Management, Wuhan University, Wuhan, Hubei, China
[b] Information Retrieval and Knowledge Mining Laboratory, Wuhan University, Wuhan, Hubei, China
[c] Department of Information Management, Peking University, Beijing, China

ARTICLE INFO

Keywords:
Scientific collaboration
Academic career
Career stages
Ego-centric networks
Scientist's role

ABSTRACT

In scientific research collaboration, researchers collaborate with different scholars throughout their career stages. Researchers at different career stages may play various roles in science teams. This paper focuses only on the researchers' roles in their respective science teams, defined as "relative roles" here, rather than comparing roles among researchers. As the academic age of researchers is increasing, researchers formulate the positive (e.g., junior-peer-senior) or negative growth trajectories (e.g., peer-peer-junior) of the relative roles throughout their career stages, which are defined as the "relative role growth patterns". Further, these growth patterns can be divided into different "relative role growth types" according to their several common characteristics. By constructing ego-centric networks of researchers based on academic age, this paper investigates the changing relative roles of researchers at different career stages and relative role growth types summarized from multiple growth patterns, and then analyzes the collaborative ability (i.e., collaboration frequencies and number of collaborators) and research performance under different growth types. In addition, we also discuss the influence of collaborators on the formation of relative role growth types. We find 13 relative role growth patterns and summarize them into four growth types. The four growth types have diverse collaborative ability and research performance. Different collaborators have different effects on the formation of growth types of researchers. Collaborators who have high productivity and high citations per paper have a positive influence on researchers' growth and vice versa. This study could be useful for researchers planning career development, also for policymakers and university administrators.

1. Introduction

There have been many studies that prove that scientific research collaboration is beneficial to research productivity and impact (Beaver & Rosen, 1978; He, Geng, & Campbell-Hunt, 2009; Lee & Bozeman, 2005), which has implications for and influence on academic career development. Research collaborators vary with a researcher's career stage, so the members of a science team are constantly changing, leading to probable changes of researchers' roles in science teams at various career stages. Therefore, the roles in this paper are only relative to the researchers' respective science teams, and are defined as "relative roles" here, rather than comparing roles among researchers. In addition, different roles in the science team have different levels of importance. For example, the academic

* Corresponding author at: School of Information Management, Wuhan University, Wuhan, Hubei, China.
E-mail addresses: weilu@whu.edu.cn (W. Lu), yanren421@whu.edu.cn (Y. Ren), yonghuang1991@whu.edu.cn (Y. Huang), buyi@pku.edu.cn (Y. Bu), john_love@whu.edu.cn (Y. Zhang).

[1] Wei Lu and Yan Ren contributed equally to this work.

https://doi.org/10.1016/j.joi.2021.101207
Received 15 February 2021; Received in revised form 20 July 2021; Accepted 18 August 2021
Available online 31 August 2021

elites who have large grants, such as principal investigators (PIs), are usually senior researchers in charge of a research project or the main manager of a scientific team, playing an important role in promoting the progress of the subject field (Kastrin et al, 2018; Melkers & Xiao, 2012). Generally, their scientific productivity, influence, number of collaborators and other scientific research performance indicators are better than those of ordinary researchers (Bozeman & Corley, 2004; De Solla Price, 1963; Feeney & Welch, 2014; Yin & Zhi, 2017), such as non-PIs. What is more, the relative roles of researchers at different career stages in turn form the relative role growth trajectory of their entire career, including positive growth trajectories (e.g., junior-peer-senior) and negative growth trajectories (e.g., peer-peer-junior). We define these growth trajectories as "relative role growth patterns" in this paper. According to several common characteristics of these growth patterns, they can be further divided into different "relative role growth types". This paper explores the relative roles (importance) of researchers at different stages in science teams and the relative role growth types of the researchers' whole academic careers, which is helpful to formulate better career development plans, help governments and universities make more reasonable scientific research policies, cultivate principal investigators, and promote the development of disciplines.

Collaboration networks have already been widely used in related research on scientific collaboration (Drożdż et al., 2017; Lande et al., 2020). As for the scientist's role/position (Glänzel, 2014) in scientific collaboration, most network studies focus on the network properties of role/position (Abbasi, Chung, & Hossain., 2012; Kong et al., 2019) and research activities (Kastrin et al., 2018; Yoshikane et al., 2009; Yoshikane, Nozawa, & Tsuji, 2006). In addition to the above literature, there are only a few studies that use ego-centric networks to explore the collaboration patterns in academic career stages (Glänzel, 2014; Wang et al., 2017) from the individual level directly. However, these investigations do not focus on the dynamic evolution of patterns of individual roles at different career stages. In this work, we construct an ego-centric network for each researcher to investigate changing patterns of roles at diverse career stages. Moreover, the roles of researchers are relative among ego-centric networks constructed in this work. The role levels only exist in each central author's ego-centric network, which is different from most of the above-mentioned studies in which senior scholars are definitely older than other roles.

We are also interested in whether collaborators have an influence on the relative role growth types of researchers. Previous studies have pointed out that co-authorship with influential collaborators has a positive impact on career success (Li et al., 2019; Liénard et al., 2018; Malmgren, Ottino, & Amaral, 2010; Peterson, 2015) by analyzing the research performance of scholars. However, few studies focus on the influence of collaborators from the perspective of relative roles. To this end, we explore whether the different co-authorship in different career stages has an impact on the formation of diverse relative role growth types.

Based on the widely used academic data set "American Physical Society (APS)" from 1998 to 2016, this paper proposes a new method to calculate the relative role of researchers and comprehensively analyze the relative role growth types summarized from multiple growth patterns, collaboration ability (i.e., collaboration frequencies and number of collaborators), research performance, and the influence of collaborators on the formation of relative role growth types using ego-centric networks. First, researchers are divided into three types of relative roles, namely junior, peer, and senior, to analyze relative role growth patterns which can be further divided into diverse growth types. Then, we analyze the collaborative ability based on collaboration frequencies and number of collaborators under different relative role growth types. Moreover, we discuss the research performance based on paper number and citations under different relative role growth types. Finally, we locate all researchers based on the research performance of their collaborators at three career stages to explore whether the collaborators have an influence on the formation of the relative role growth types.

Three main questions are laid out in this work:

RQ1. What are the relative role growth patterns and growth types of the scientific researchers?
RQ2. What is the collaboration ability and scientific research performance under different relative role growth types?
RQ3. Do the collaborators in different career stages have an impact on the formation of diverse relative role growth types?

2. Related work

2.1. Ego-centric network in career analysis

With the great emergence of various scholars' research evaluation indicators, the research focus on the collaboration patterns using co-authorship network analysis has gradually shifted from a macro perspective to a micro perspective. Many related network studies have been carried out using ego-centric networks. The ego-centric network is part of the whole network structure, which can be searched through a breadth-first search via a given node. A network with the above characteristics can be easily analyzed across the whole network and can be used to draw statistically important conclusions about the entire population (Everett & Borgatti, 2005).

Since career analysis at the individual level became possible through quantitative science studies, most scholars have begun to zoom in on the two categories using ego-centric networks. In the first category, the relations between network properties and scholarly performance are investigated. Abbasi et al. (2012) applied structural hole theory to explore how ego-centric network properties of density, efficiency, and constraint relate to research performance. They noticed that scholars with higher levels of betweenness centrality and efficient co-authorship network perform better in research. Ortega (2014) analyzed the relationship between research impact and the structural properties of ego collaboration networks. Results show that sparse ego-centric networks help central authors control collaborators and have a positive impact on research impact. Badar, Frantz and Jabeen (2016) examine the association of co-authorship network centrality and research performance. They find higher centrality has a better performance in academic research. In addition to the network properties, another category is collaboration pattern in academic research. Petersen (2015) studied research collaboration pattern from an ego-centric perspective. He found that super ties have a significant impact on productivity and citations throughout scholars' careers. Wang et al. (2017) investigate scholars' academic-age-aware collaboration patterns in Physics and Computer Science. Their results show that senior scholars tend to have more collaborators and the obvious homophily phenomenon in collaborations.

Use of the ego-centric network in career analysis provides new insights into its application in studies of a scientist's role/position. However, these studies primarily focus on the network structure and collaboration patterns, failing to address the scholars' own career development. Therefore, the present work will concentrate on the researchers' role growth patterns across their career.

2.2. Role classification and career development in career analysis

As scholars turn their research focus to the individual scientists, research into the scientist's role/position (Glänzel, 2014) has been widely carried out in terms of career analysis, which is always inseparable from the role classification. There are several classification criteria such as academic age, rank order, and position in a project. For example, according to the length of academic age, Wang et al. (2017) divided the roles into beginning scholars, junior scholars, and senior scholars. According to the rank order of the publication, Milojević, Radicchi and Walsh (2018) divided the authors' knowledge production roles into lead authors and supporting authors, and divided them into transients, dropouts, and full career scientists according to the dimension of each researcher's publication time. Furthermore, according to whether researchers take the lead in scientific research projects, Feeney and Welch (2014) divided the roles into PIs and co-principal investigators (co-PIs) according to their lab affiliations and grants. Nevertheless, the classification criteria of roles in most of the above-mentioned studies are absolute. For example, for the classification criterion of academic age, the role will become more and more senior (influential) with the increase of academic age (senior>peer>junior). In this work, there are no levels among the relative roles of the central authors. We focus only on the researchers' roles in their respective science teams, defined as "relative roles" here, rather than comparing roles among researchers, which is different from the role studied before.

An author's career development remains the primary issue in career analysis. Many studies have found that diverse factors can lead to the success or failure of academic careers. Wang et al. (2019) found that a beginning-career setback has a great negative impact on career success from the perspective of funding. Feeney and Welch (2014) demonstrated that whether a researcher is a PI does have an impact on scientific research output: the scientific output and grants of PI researchers are always higher than those of co-PIs. Kong et al. (2019) studied the influence of collaboration on the position of researchers and found that cooperation can increase the influence of authors. Sugimoto et al. (2016) carried out a study of scholarly communication behavior. They found that the role of scholars is related to research output, collaboration, and influence. Specifically, the productivity of scholars rises sharply until they are promoted to associate professors and then it remains stable, and the older a scholar is, the more they collaborate. Kong et al. (2020) found that author-centered and article-centered factors have a great impact on the career success of scholars, indicating that collaborators matter in a scholar's career.

From the above-mentioned studies, on the one hand, few studies use scholars' influencing indicators, such as h-index, as the criteria for role classification and roles in those studies are not relative. On the other hand, most investigations of career development in role studies concentrate on the factors influencing research performance. However, few studies have explored the career development from the perspective of role evolution. Therefore, we focus on the relative role evolution phenomenon of researchers in Physics in this work.

The remainder of this paper proceeds as follows. We detail the data processing and methods utilized for our analysis. We next provide and present our findings. In addition, we discuss the results of the paper and conclude the study with a summary. Finally, we point out the implications and limitations of this study, and suggest directions for future research.

3. Methodology

3.1. Data

The dataset used in this work is obtained from "American Physical Society (APS)", which contains a number of journals in the physical review series which have a high reputation among researchers in the global physics community and related disciplines. The dataset includes 491,197 authors and 596,786 collaboration pairs between 1998 and 2016. Author names are disambiguated according to Bai, Zhang and Lee (2019), in which a unified probabilistic framework is implemented along with both content- and citation-based information and three steps are included: First, we judge if the authors' last names are identical; second, we judge if the authors' first names or initials are identical; and, finally, we make sure that one of the three following conditions is true: whether the authors cited each other at least once or have one co-author or similar affiliation at least.

In this work, we use the h-index (Hirsh, 2005) as the proxy to calculate the relative roles of each researcher in ego-centric networks and explore their relative role growth patterns. This index has already been widely accepted by academia as an indicator to evaluate a scientist's lifetime achievement regardless of its several disadvantages. It can identify influential (Cronin & Meho, 2006) or senior scholars (Li & Gillet, 2013) in a certain research area, coauthoring with whom can give juniors consistent advantages in their academic career (Li et al., 2019). Therefore, the h-index is an appropriate indicator to represent the research capacity of researchers in order to decide the node size when constructing ego-centric networks in this work.

We select researchers who have written at least 50 articles and whose academic age is not less than 30 years, providing a final dataset of 1315 unique central researchers. The 1315 scientists only refer to the central nodes, which are called central researchers

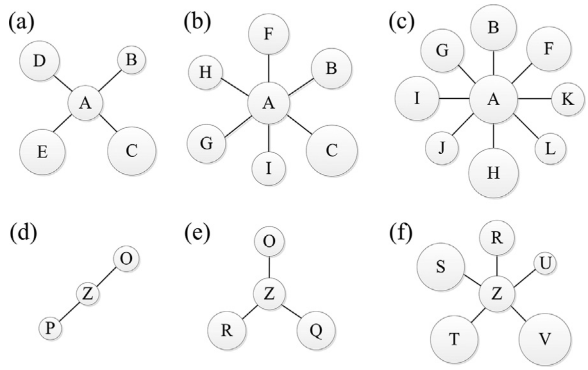

Fig. 1. (a) An ego-centric network of the central researcher "A" at a certain academic age at the beginning career stage. (b) An ego-centric network of the central researcher "A" at a certain academic age at the mid-career stage. (c) An ego-centric network of the central researcher "A" at a certain academic age at the late career stage. (d) An ego-centric network of the central researcher "Z" at a certain academic age at the beginning career stage. (e) An ego-centric network of the central researcher "Z" at a certain academic age at the mid-career stage. (f) An ego-centric network of the central researcher "Z" at a certain academic age at the late career stage.

below. In this article, we focus the 1315 central researchers and the collaborators of them. Academic age is calculated by the year when an author published his or her first article and the year in which they published their last article according to all their past publications. The h-index is calculated according to each researcher's publication and citation counts recorded in the dataset.

3.2. Network construction

To study the changing relative roles of researchers at different career stages in the field of physics, all-academic-age undirected ego-centric networks of 1315 central researchers are constructed after author disambiguation.

In social networks, an ego-centric network refers to a network that includes a central node with other nodes and edges connecting to it. Such networks often focus on the nature of individual nodes rather than on the entire network. In ego-centric networks in this work, the nodes represent researchers, and the edges represent the collaboration relationship between researchers. In order to visually show the changes of the researcher's relative roles in the ego-centric network with academic age, the size of the node in an ego-centric network depends on the researcher's h-index.

Fig. 1 shows examples of the ego-centric networks of the central researchers "A" and "Z". In the ego-centric networks of researchers "A" and "Z", the node represents the researcher, the node size represents the researcher's h-index, the letter in the node represents the researcher's name, and the edge between the nodes represents the collaboration relationship between the researchers. Supposing that the career is divided into three stages, Figs. 1(a), 1(b) and 1(c) respectively represent the ego-centric network of the central researcher "A" at three different academic ages and thus three different career stages. Figs. 1(d), 1(e) and 1(f) respectively represent the ego-centric network of the central researcher "Z" at three different academic ages and also three different career stages.

Taking researcher "A" as an example, as researcher "A" increases in age, their career stage changes as do their collaborators and collaboration frequencies. The node size reflects influence and importance of researchers in their science team to some extent. The larger the node size, the higher the influence. In Fig. 1(a), 1(b) and 1(c), the node size of researcher "A" increases during their career, and, due to the change of collaborators, the comparison of the node size of researcher "A" and the node size of collaborators reflects the varying degrees of importance of the central researcher "A" at different academic ages and different career stages in the science team.

3.3. Definition of relative role

Most of the author roles mentioned in previous studies were calculated and classified based on absolute metrics, such as academic age (Wang et al., 2017) and rank order (Milojević et al., 2018). According to these roles, from a certain point of view, the scientists who are the research objects can be directly divided into different levels. For example, according to the academic age, then the academic age of seniors must be longer than junior and peer scientists. In this work, we propose "relative role" of the central researchers in their science teams. That is to say, we focus only on the researchers' roles in their respective science teams, defined as "relative roles" here, rather than compare roles among central researchers. To be specific, there is no level among the relative roles of the central researchers. The levels of roles of the central researchers are only relative to their own collaborators in their ego-centric networks. For example, supposing the relative roles of the central researcher "A" and central researcher "Z" are respectively junior and peer in their own ego-centric networks according their collaborators (Figs. 1a and d), however, the influence of researcher "A" is greater than researcher "Z" according their node size representing by h-index.

Relative roles represent the influence of the central researchers relative to their collaborators in the ego-centric networks, and also the importance of the central researchers to their science teams. Constantly seeking importance in the team conforms to the fourth level of the Maslow model (Wahba & Bridwell, 1973) to a certain extent, and the more important roles in the team, such as

PI, seem to be able to attract academic resources. However, some researchers may take long time to reach this position, and even some researchers may not reach this position in their entire academic career. Due to the change of collaborators, the relative roles of central researchers in their science teams may also vary. The relative role of each academic career stage represents to a certain extent the researcher's career development status at this stage. Therefore, the relative roles of different stages reveal the growth pattern of a researcher in his/her career. By exploring the different growth patterns of the relative roles of researchers in their careers, this article helps researchers to formulate reasonable career plans and provides research policy suggestion for universities and governments.

3.4. Measurement calculation

3.4.1. Measurement of relative role

Regarding the calculation of the relative role of researchers in the ego-centric network, we propose a Relative role-based h-index (Rrhi) to measure each central researcher's relative role in his/her ego-centric network. Taking researcher "A" as an example, at a certain academic age, there are N nodes (j belongs to N, N≥j≥1) of collaborators of the central researcher "A" in the ego-centric network, then the Rrhi of researcher "A" is calculated as:

$$\mathrm{Rrhi(A)} = \frac{\mathrm{hi(A)}}{\sum_{j=1}^{N} \frac{hi_j}{N}}$$

where hi(A) is the h-index of the central researcher "A" at the certain academic age, $\mathrm{hi_j}$ is the h-index of each collaborator in his/her ego-centric network, and N is the number of collaborators of the central researcher "A" at the certain age. Rrhi is the ratio of the h-index of the central researcher at a certain academic age to the average h-index of the collaborators in that age. Taking researcher "A" in Fig. 1(a) as an example, $\mathrm{Rrhi(A)} = \frac{\mathrm{hi(A)}}{\frac{\mathrm{hi(B)+hi(C)+hi(D)+hi\ (E)}}{4}}$.

3.4.2. Measurement of collaborative ability

The following measures related to the collaborative ability are considered.

- Collaboration frequency: The number of co-authored publications of the two researchers. If one collaborator recurs in two or more papers of a certain researcher, we count the frequency according to the number of papers.
- Average collaboration frequency: the average collaboration frequency of researchers in a given relative role growth pattern.
- Number of collaborators: The number of collaborators the central researchers has. If one collaborator recurs in two or more papers of a certain researcher, it is counted as the same one collaborator.
- Average number of collaborators: the average number of collaborators of researchers in a given relative role growth pattern.

3.4.3. Measurement of research performance

The following measures related to the research performance are considered.

- Number of papers: number of papers included in the American Physical Society (APS) database.
- Average number of papers: the average number of papers of researchers in a given relative role growth pattern.
- Total citations: total citations are received by papers in APS database. The citations are counted from publication year to 2016.
- Average total citations: the average total number of citations of researchers in a given relative role growth pattern.
- Citations per paper: average citations of each paper of a given researcher.
- Average citations per paper: the average citations per paper of researchers in a given relative role growth pattern.

3.5. Classification of relative roles and academic career stages

The relative roles of researchers in this paper are divided into three types: Juniors, Peers and Seniors (Glänzel, 2014). Juniors may have the least research experience and may seldom play a decisive role in the process of collaboration in their own ego co-authorship networks. Peers approximatively equal the status of the neighbors in their own ego-centric networks, and usually play a similar role to their collaborators. Seniors refers to the dominant role in the process of collaboration, and they usually play the role of planning, supervising, and guiding the overall situation in collaboration in their own ego-centric networks.

In order to divide the researchers into different relative roles and divide each researcher's academic career into different stages, we first calculate the number of researchers of Rrhi in the APS database, and rank Rrhi based on the proportion of researchers. Specifically, the researchers whose number of published articles is more than 50 and whose academic age is between 30 and 60 in the database are included in the statistics. Fig. 2 shows the proportion of researchers corresponding to different values of Rrhi.

In Fig. 2, we first calculate the proportion of the researchers corresponding to different values of Rrhi, and then obtain the two x-axis values of Rrhi according to the corresponding image intersection point based on this average value of y-axis, and then divide Rrhi into three levels for the following specific analysis. From the perspective of the proportion of researchers corresponding to the values of Rrhi in a certain field shown in Fig. 2, the number of researchers is different for diverse Rrhi, and the number of researchers rises rapidly as the Rrhi increases and then gradually decreases. Specifically, when Rrhi<1, the proportion of researchers rises faster; when 1≤Rrhi<4, the proportion of researchers is in a range where it is relatively high; when Rrhi≥4, the proportion gradually decreases and presents a "long tail" distribution. According to this phenomenon, the relative roles of researchers can be divided into three groups: juniors (Rrhi<1), peers (1≤Rrhi<4), and seniors (Rrhi≥4).

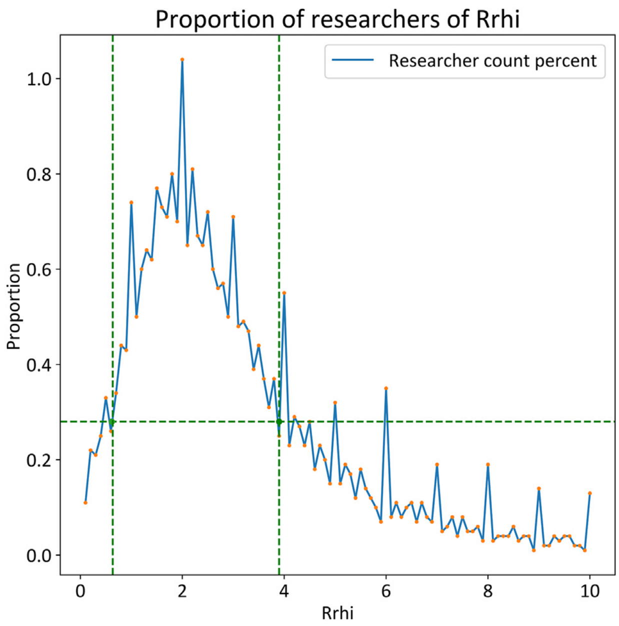

Fig. 2. Distribution of researchers corresponding to the values of Rrhi.

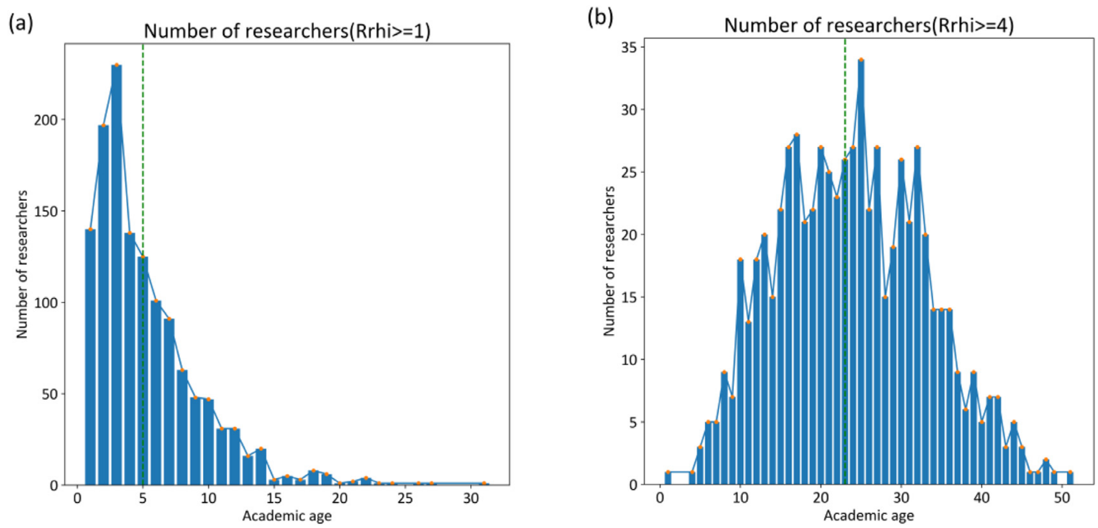

Fig. 3. (a) Central researchers' Rrhi reaches 1 for the first time. (b) Central researchers' Rrhi reaches 4 for the first time.

According to Rrhi's level, we calculate the average academic age of all researchers in the database when their values of Rrhi reach 1 and 4 for the first time, as shown in Figs. 3(a) and 3(b). When a researcher's value of Rrhi reaches 1 for the first time, it means that the researcher starts to have a similar influence to their collaborators; when a researcher's value of Rrhi reaches 4 for the first time, it means that the researcher starts to have a higher influence than most of his collaborators, playing a more important role in the collaboration. According to the above calculation result of the average academic age, we divide the academic career evolution process of researchers in this article into three stages: Beginning career stage (academic age≤5), mid-career stage (5<academic age≤23) and late career stage (academic age>23).

The relative role of a researcher at a certain career stage depends on the average value of Rrhi at the certain career stage. For example, as the average Rrhi of researcher "A" at the beginning career stage is 0.6 (Rrhi<1), then the relative role of researcher "A" at this stage is junior.

All data processing and image visualization in this work are implemented using Python tools, and the network construction is implemented using the NetworkX software package in Python.

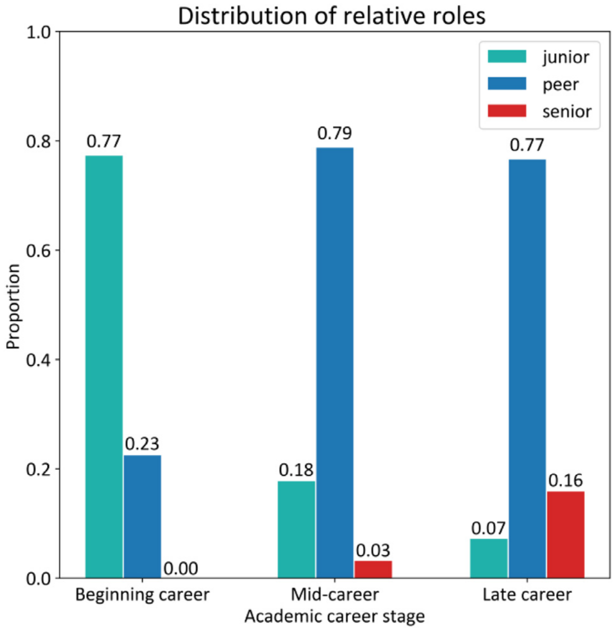

Fig. 4. Distribution of the proportion of relative roles respectively at the beginning career stage, the mid-career stage and the late career stage.

Table 1
Statistics of the researchers' relative role growth patterns.

Serial Number	J growth	J effective	Proportion	P growth	P effective	Proportion
1	JJJ	√	1.1%	PPP	√	7.7%
2	JJP	√	16%	PPJ	√	1.2%
3	JJS	√	0.3%	PPS	√	1.0%
4	JPP	√	62.7%	PJJ	×	-
5	JPS	√	5.5%	PJP	√	0.2%
6	JPJ	√	3.7%	PJS	×	-
7	JSS	√	0.2%	PSS	√	0.2%
8	JSP	×	-	PSJ	×	-
9	JSJ	×	-	PSP	√	0.2%

4. Result

4.1. Relative role growth patterns and growth types

In order to further study the researchers' relative role growth patterns, we first analyze the distribution of researchers' relative roles at three career stages. We count the proportion of the relative roles at each of the three career stages, as shown in Fig. 4.

From the overall situation, the relative roles with the largest proportion at the three career stages are respectively junior, peer, and peer. Seniors account for very little at the three career stages. The proportion of juniors rapidly declines from 77% to 18% at the mid-career stage and slowly declines to 7% at the late career stage. The proportion of peers rapidly increases from 23% to 79% at the mid-career stage and slightly declines to 77% at the late career stage. Seniors do not appear at the beginning stage and slowly increase from 3% to 16% at the late career stage.

Based on the work above, we find 13 relative role growth patterns of 1315 researchers after they have experienced the three stages of beginning career, mid-career and late career (their academic ages are all over 30). There are 3^3=27 types of evolutionary combinations of the three roles in theory. However, there are only 13 effective patterns are found in practice. Table 1 summarizes the 13 relative role growth patterns and the corresponding proportions.

According to the proportions, they are JPP (62.7%), JJP (16%), PPP (7.7%), JPS (5.5%), JPJ (3.7%), PPJ (1.2%), JJJ (1.1%), PPS (1.0%), JJS (0.3%), PJP (0.2%), JSS (0.2%), PSS (0.2%), and PSP (0.2%). Table 1 illustrates the multiple role growth patterns of the researchers. From the perspective of role development at the career stage, these patterns are further divided into four growth types: rapid growth (0.6%), general growth (85.5%), stagnant growth (8.8%), and regressive growth (5.1%). Furthermore, we randomly select a researcher's ego-centric networks from each of the three academic career stages respectively to show the dynamic network structure of some patterns in these four growth types.

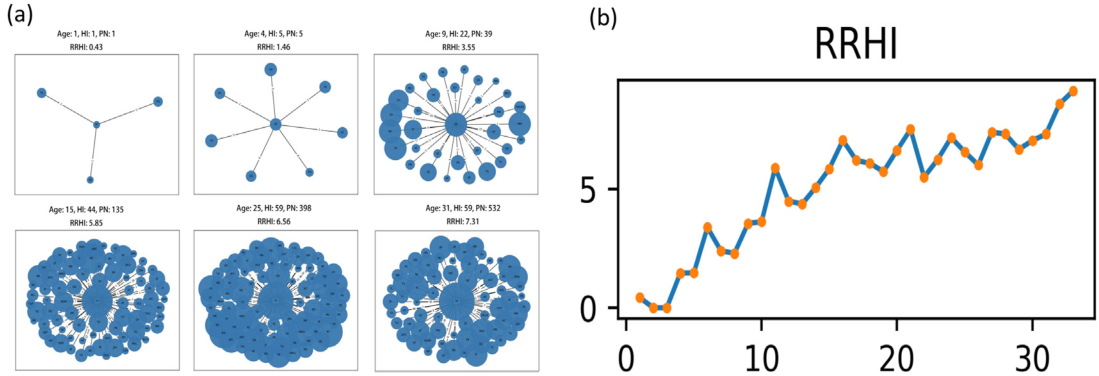

Fig. 5. (a) Ego-centric networks of an example researcher “YT” of the JSS pattern. (b) YT’s Rrhi curve with academic age.

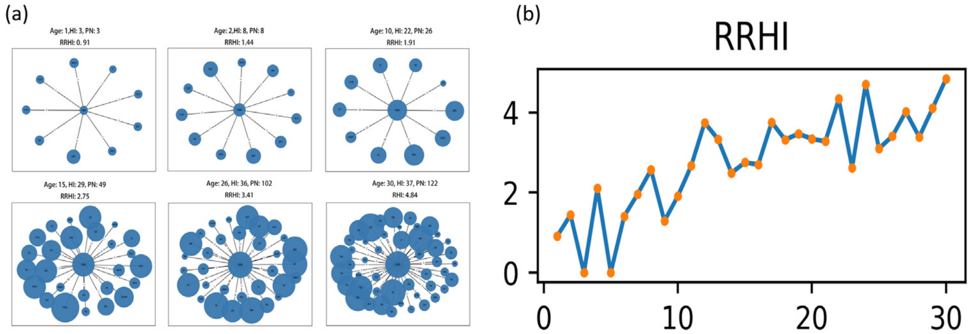

Fig. 6. (a) Ego-centric networks of an example researcher “DAB” of the JPS pattern. (b) DAB’s Rrhi curve with academic age.

4.1.1. *Rapid growth type*

The rapid growth type refers to the patterns in which researchers directly convert into the relative role of senior at the mid-career stage, which comprises the following three patterns: JSS (0.2%), PSS (0.2%), and PSP (0.2%).

It is found that there are two patterns, JSS and PSS, with senior roles at two career stages, defined as “double S” patterns. The proportion of these two patterns is 0.2% respectively, which is extremely small. PSP is a special pattern with two characteristics, rapid growth and regressive growth. According to the analysis above, the proportion of seniors is relatively small, as they only appear at the mid- and late career stages in small proportions. Therefore, we put PSP in the rapid growth type, not in the regressive growth type. However, whether the collaborative ability and research performance of PSP is similar to the other two patterns in this type is worthy of further discussion. Fig. 5 shows an example of a researcher’s dynamic ego-centric networks.

4.1.2. *General growth type*

General growth type refers to the patterns in which researchers realize the forward growth and transformation of the relative role at the mid- or late career stages. It contains the following five patterns: JPP (62.7%), JJP (16%), JPS (5.5%), PPS (1.0%) and JJS (0.3%).

This type of role growth pattern accounts for the largest proportion (85.6%) in the four growth types, indicating that most researchers develop their roles in their careers. It is found that the pattern of JPP accounts for 62.7%, which is the largest proportion in the 13 patterns, indicating that most researchers are in the relative role of junior in the beginning career stage and have a role promotion at the mid-career stage. We also find that this type contains one of the most normally developed patterns, JPS. However, this pattern accounts for only 5.5% of the 13 patterns, indicating that the researchers who experience the relative roles of junior, peer, and senior at the three career stages are a minority. Both PPS and JJS are promoted to the role of senior at the late career stage, which accounts for a small proportion (1.3%) compared to other patterns in this type. Fig. 6 shows the ego-centric networks of an example researcher in this type.

4.1.3. *Stagnant growth type*

The stagnant growth type refers to the patterns in which researchers are in the same relative role at all three career stages. It comprises the following two patterns: PPP (7.7%) and JJJ (1.1%).

The proportion of PPP ranks third among the 13 patterns, which shows that a certain proportion of researchers are in the relative role of peers in the beginning career stage, and they maintain the same role across the three career stages, collaborating with scholars who have a similar influence to themselves at the three career stages. JJJ accounts for a small proportion, indicating that few

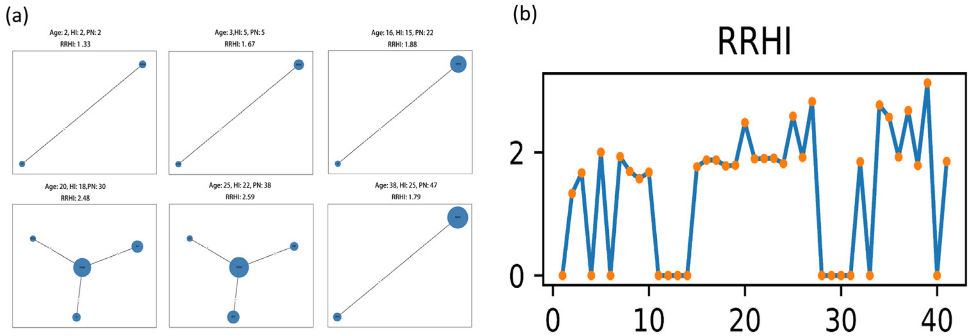

Fig. 7. (a) Ego-centric networks of an example researcher "RMW" of the PPP pattern (Robert M. Wald #A1#). (b) RMW's Rrhi curve with academic age.

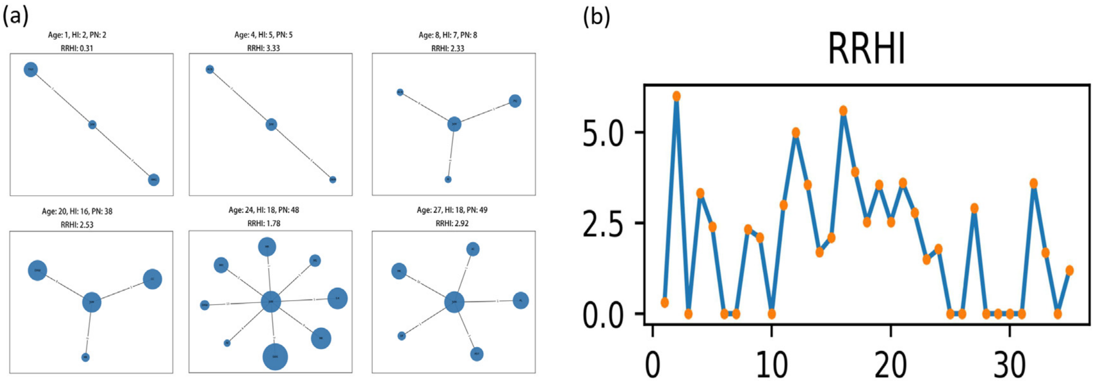

Fig. 8. (a) Ego-centric networks of an example researcher "JVM" of the PPJ pattern. (b) JVM's Rrhi curve with academic age.

researchers maintain a beardless role in all three career stages. Fig. 7 shows the ego-centric networks of an example researcher in this type.

4.1.4. Regressive growth type

The regressive growth type refers to patterns in which researchers' relative roles regress at the mid- or late career stages. It comprises the following three patterns: JPJ(3.7%), PPJ(1.2%), and PJP(0.2%).

The total proportion of this type is small, adding up to 5.1%, indicating that most researchers pursue role improvement in the science team, and few researchers experience retreat in their academic careers. In this type, however, there exists possibility that some senior scientists may collaborate with more senior (higher h-index) people in their late career. Considering this situation, we take the pattern JPJ as an example and investigate the ego-centered networks of all 48 scientists under the pattern JPJ and their corresponding Rrhi values for each academic year. We find that few scientists collaborate with authors with greater influence than their own in the late career (higher influence authors accounting for the majority in the ego-centered network). The statistical results show that only five scientists collaborate with higher influential authors in 1-2 years in their late career, so this situation has little effect on the overall late career stage of these scientists in the present work and we cannot concentrate on this phenomenon by using the present dataset. Fig. 8 shows the ego-centric networks of an example researcher in this type.

According to the rise and regression of the relative role, we divide the four types into two categories: positive growth and negative growth. It can be seen from the ego-centric networks of the four growth types that the changes of relative role have two circumstances: Researchers' own growth; the change of the collaborator or science team. Collaborative capability and scientific research performance of the four growth types need to be further studied.

4.2. Collaborative ability of diverse growth types

We use average collaboration frequencies and average number of collaborators as indicators to measure collaborative ability. As for the average collaboration frequencies, if one collaborator recurs in two or more papers of a certain researcher, we count the frequency according to the number of papers. As for the average number of collaborators, if one collaborator recurs in two or more papers of a certain researcher, it is counted as the same one collaborator. 0 shows average collaboration frequencies and average number of collaborators in the four growth types. Fig. 10 shows the rank of patterns in the four growth types by average collaboration frequencies and average number of collaborators.

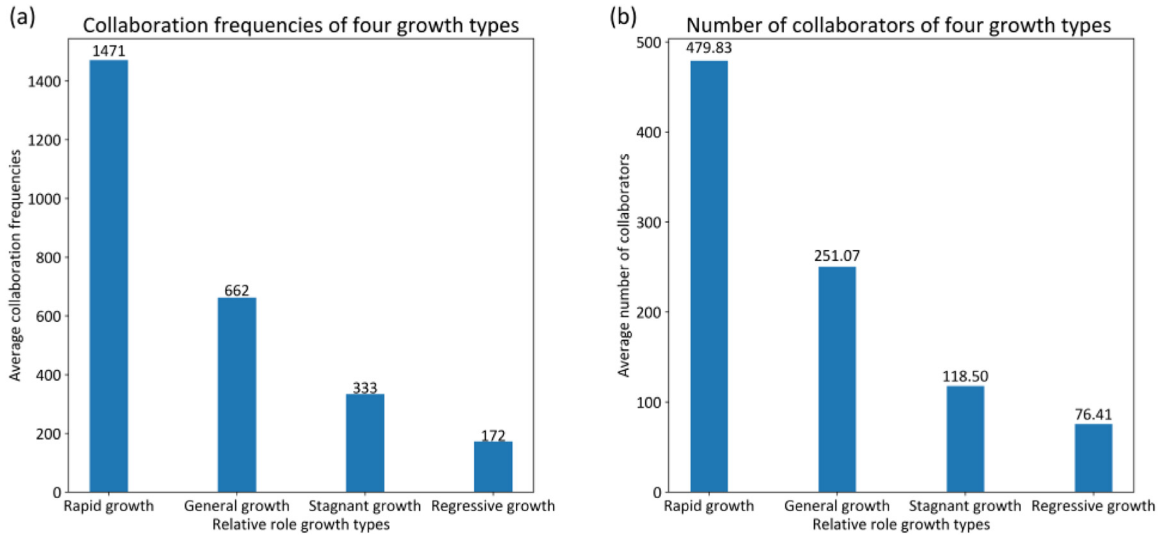

Fig. 9. (a) Ranks of the four growth types by average collaboration frequencies. (b) Ranks of the four growth types by average number of collaborators.

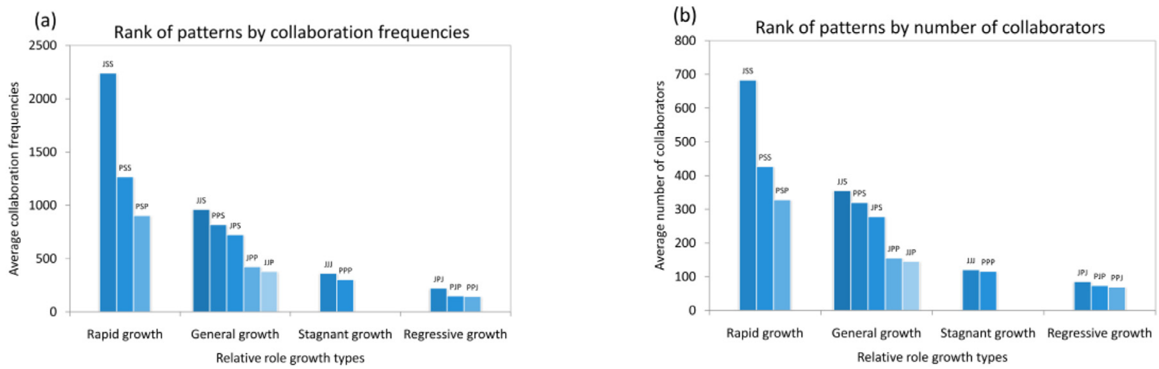

Fig. 10. (a) Rank of patterns by average collaboration frequencies in the four growth types (b) Rank of patterns by average number of collaborators in the four growth types.

As can be seen from Figs. 9 and 10, the ranks of average collaboration frequencies are the same as the ranks of average number of collaborators in the four growth types. The same situation also appears in the ranks of patterns by the average collaboration frequencies and average number of collaborators in the four growth types.

It can be clearly seen that the rapid growth type ranks at the top in terms of collaboration frequencies and average number of collaborators, followed by the general growth type, the stagnant growth type, and the regressive growth type. The two indicators of collaborative ability of the rapid growth type far exceed those of the other types, especially the collaboration frequencies. Among the three patterns in this type, the JSS and PSS patterns with two “seniors” in the career stages perform best. PSP as a pattern with a characteristic of regression, but its two indicators are higher than that of most of other patterns. The general growth type ranks second and there are also some patterns that perform well, such as JJS, PPS, and JPS. Both the stagnant growth type and the regressive growth type are far behind the general growth type, with the stagnant growth type ranking higher than the regressive growth type.

4.3. Research performance of diverse growth types

We investigate the research performance of the four growth types by productivity and citations. We use number of papers as an indicator to measure research productivity, and use total citations and citations per paper as indicators to measure research impact in this work. Fig. 11 shows the average number of papers, average total citations, and average citations per paper in the four growth types. Fig. 12 shows ranks of patterns by average number of papers, average total citations, and average citations per paper in the four growth types.

As for number of papers and total citations, the rapid growth type performs the best in the four types, followed by general growth type, stagnant growth type, and regressive growth type. As for the citations per paper, things are different. The gap between the four growth types narrows, and the rapid growth type is no longer ranked first, but is replaced by the general growth type. The regressive

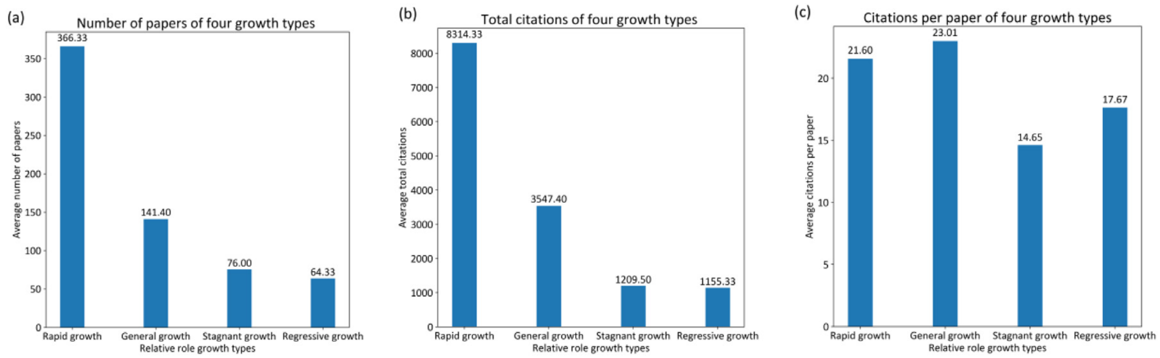

Fig. 11. (a) Ranks of the four growth types by average number of papers. (b) Ranks of the four growth types by average total citations. (c) Ranks of the four growth types by average citations per paper.

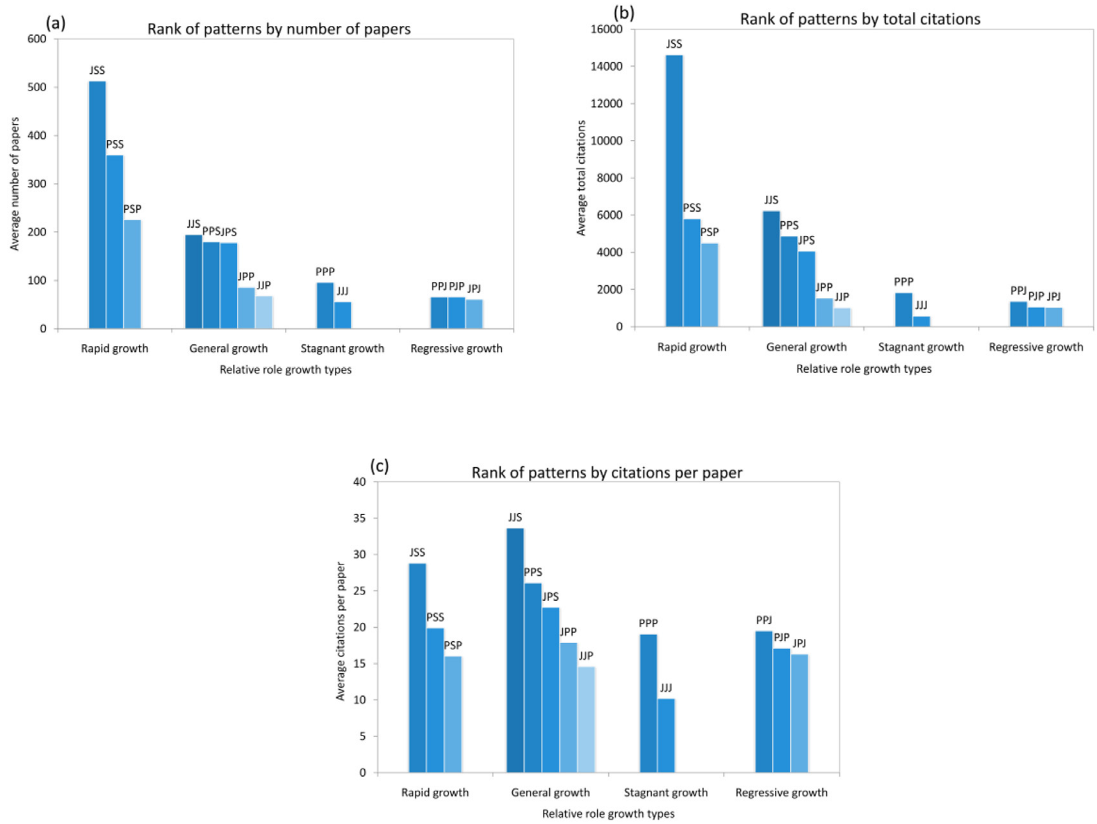

Fig. 12. (a) Ranks of patterns by average number of papers in the four growth types. (b) Ranks of patterns by average total citations in the four growth types. (c) Ranks of patterns by average citations per paper in the four growth types.

growth type is not ranked at the end, but is higher than the stagnant growth type. However, there exist some common phenomena of the ranking of patterns: On the one hand, the research performance of patterns with the relative role of "senior" is always better than that of other patterns in the same growth type, such as JSS, JJS; on the other hand, the ranks of patterns are exactly the same among the growth types, besides the patterns with "senior", which is the same as the pattern ranking of collaborative ability. It is worth noting that research performance should be comprehensively considered by the three indicators. Therefore, the research performance of the rapid growth type is ranked top, followed by the general growth type, stagnant growth type, and regressive growth type. However, the gap between the stagnant growth type and regressive growth type is small.

It is found that the research performance of JSS and PSS in the rapid growth type maintains its dominant advantage in all 13 growth patterns, except for the citations per paper. PSP perform better than most of other patterns, as the collaborative ability shows, indicating that the regression of this pattern not certainly represent an obvious setback at the late academic career stage. Like the

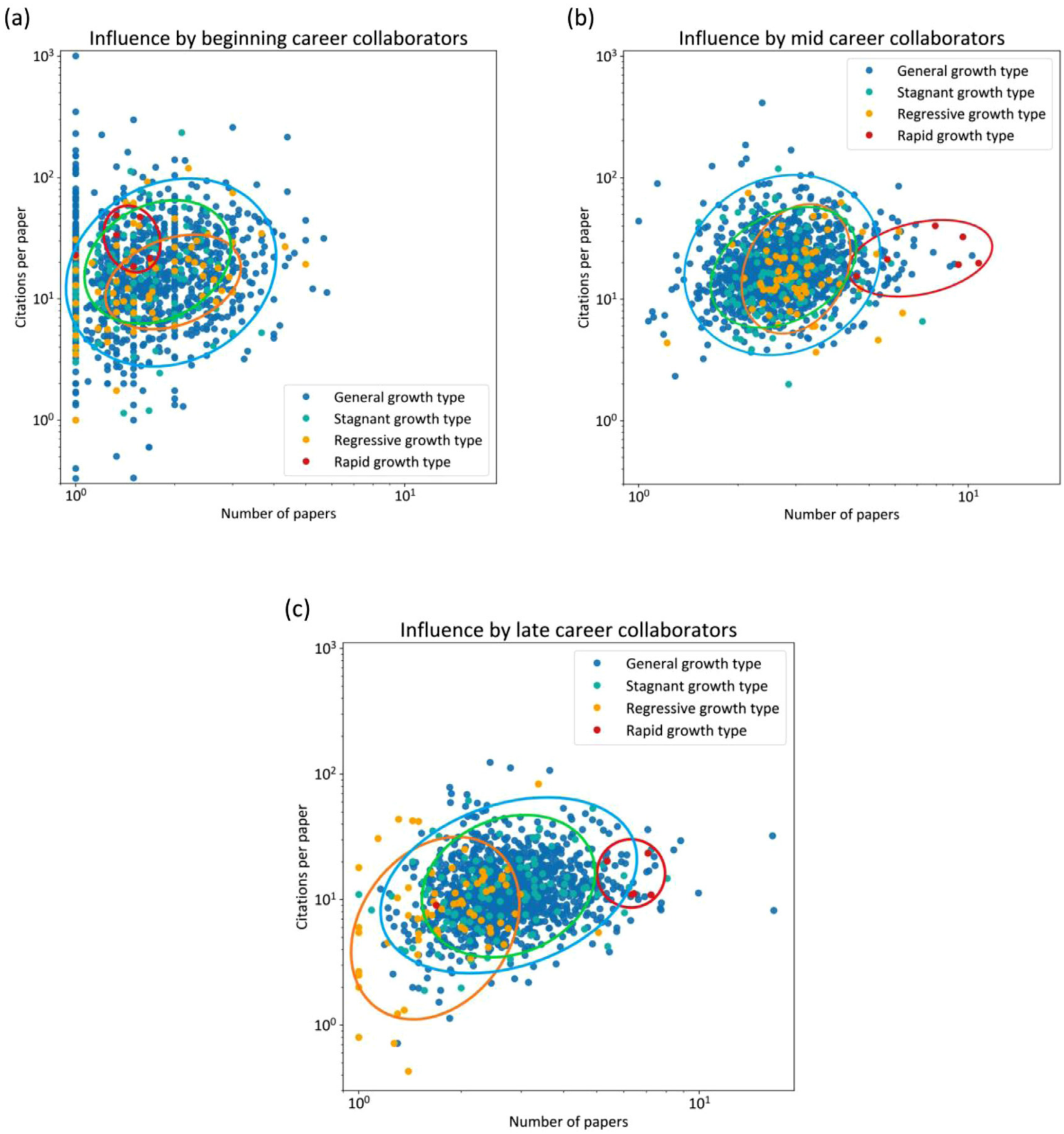

Fig. 13. (a) The locations of 1315 central researchers of the four growth types in the beginning career stage by their collaborators. (b) The locations of 1315 central researchers of the four growth types in the mid-career stage by their collaborators. (c) The locations of 1315 central researchers of the four growth types in the late career stage by their collaborators.

two indicators of collaborative ability, the patterns of JJS, JPS, and PPS with the relative role of senior perform better in terms of productivity and impact in the general growth type, while other patterns, such as JJP, do not perform well in this type. In the stagnant growth type, the total citations and number of papers of JJJ are always lower than PPP (even if the collaboration ability of JJJ is better than that of PPP).

4.4. Influence of collaborators on diverse growth types

In this section, we explore the impact of collaborators on the formation of the researchers' growth types at different stages. We use the number of papers and citations per paper of collaborators at each career stage to locate the central author in the coordinate axis after a scale log, in which the x-axis represents number of papers and the y-axis represents citations per paper. Then, we circle the areas where researchers of different growth types are located. Fig. 13 shows the 1315 researchers of the four types in the three career stages located by their collaborators.

As we can see from Fig. 13, the researchers' locations in the four growth types are different at the three career stages due to the different research performances of their collaborators.

The rapid growth type group is the smallest, but the relative position shift in the coordinate axis is the most obvious. Most researchers of this type move to areas of high number of papers and high citations per paper at the mid- and late career stages. The collaborators of the rapid growth type perform best among the four growth types overall. The researchers of the rapid growth type collaborate with the most productive and influential collaborators at the mid- and late career stages, which enables researchers to reach the relative role of senior at the mid-career stage and maintain the advantage at the late career stage.

As for the general growth type, the collaborators of this type are more scattered and the group of this type is the largest of the four growth types. Therefore, although the number of papers and the citations per paper increase slightly in the mid- and late career stages, there are a few researchers located at high impact and productivity areas around the edge of the circle or outside the circle. The research productivity and impact of the collaborators slightly increase, while the relative roles of the researchers in the second or third stages also develop.

As for the stagnant growth type, collaborators of this type have a slight productivity increase at the mid-career stage and stop increasing at the late stage, with little change of citations per paper across their careers. Researchers of the stagnant growth type collaborate with scholars who have no change of research performance in the late career stage, leading to the researchers staying in the same roles. As for the regressive growth type, the number of papers and the citations per paper improve at the mid-career stage and then retreat to a low productivity and impact area that is basically the same as that of the beginning career stage.

Researchers in the regressive growth type experience a significant decline in the number of papers and citations per paper in the late career stage. Therefore, the relative role of most researchers of this type regresses at the late career stage (several researchers' relative roles regress at the mid-career stage).

5. Discussion and conclusion

In this work, we investigate the relative role growth types of researchers in the field of physics in their academic careers and the influence of collaborators on the formation of relative role growth types. On the one hand, the career growth pattern of researchers is not the same, but has multiple relative role growth types. We find that the four relative role growth types of researchers can be further divided into two categories: Positive growth and negative growth. Positive growth types tend to have better research performance, as well as collaborative ability, especially the rapid growth type. On the other hand, the researcher's collaborators at each stage have an impact on the relative role growth type. Working with collaborators who have good performance tends to promote the formation of positive relative role growth category and vice versa.

Career development is a common problem faced by researchers in the process of scientific research. The role or position of the researcher in the science team is closely connected to career development mainly reflected by research performance and other factors such as collaborations (Sugimoto et al., 2016). Unlike the numerous extant works that study the career success by defining the scholars' roles according to the absolute classification criteria, such as the length of academic age or paper author rank (Milojević et al., 2018; Wang et al., 2017), we propose relative role using the relative index "Rrhi" to dynamically calculate the level of central researchers. The huge proportion of junior at the beginning career stage indicates the "Pareto's principle" of academia. That is to say, few scientists can be the core of the team when they first enter the academia. The proportion of different relative roles changes over time, indicating that the relative growth patterns of researchers are not always the same. The proportion of the general growth type is the largest among the four types, followed by the stagnant growth type, the regressive growth type and the rapid growth type. It indicates that there are only few people can reach and maintain the senior position in the team meanwhile at the mid and late career stages. Most researchers develop in the ordinary way in their careers and a small number of roles of researchers are static or regress, indicating that the vast majority of researchers are still seeking a position improvement in their careers.

By analyzing relative roles, we find that the research performance of researchers under the rapid growth type is higher than those of other types of researchers, especially the regressive growth type. Moreover, the research performance of researchers under the patterns with the "senior" role is higher than those of other patterns in each relative role growth type. This corresponds to Sugimoto et al.'s (2016) finding of the relationship between rank order and productivity where researchers in higher ranks are more productive. In addition to productivity, we also explore the relationship between influence and researcher's role where high level of relative role is more influential. Besides, we discover that the relationship between the collaborative ability and the researcher's role is the same as that of the research performance (Bordons et al., 2015), and researchers who own stronger collaborative ability have better research performance. Similar research was also reported by Li et al. (2018) in a study of institutions of interinstitutional scientific collaboration networks in materials science.

The above results reflect the Matthew effect in academia which suggests that scientists who have previously been successful are more likely to succeed again. That is to say, the more credits, grants, and accumulations researchers have, the easier they can attract academic resources and collaborators (Bol, de Vaan, & van de Rijt, 2018; Merton, 1957; Tol, 2013), and their productivity and influence are higher than those of ordinary researchers (Feeney & Welch, 2014), which is consistent with our conclusions. Therefore, researchers should be aware of their relative roles in science teams, and should strive to become an important force (such as PI) in the team as soon as possible. In addition, researchers would be better to try to maintain their advantages and important roles or position in the team. Universities and governments could consider focusing more policies and opportunities on young members of the academic elite to enable them to explore their team leadership skills and thus avoid the negative effects of the Matthew effect in academia.

Considering the benefits of positive growth categories for career development, we further investigate the influence of collaborators on the formation of relative role growth types. Our results show that the productivity and citations of collaborators have an influence on the researchers' locations. Further, collaborators with different research performances have an impact on the formation of relative role growth types. In a certain career stage, working with collaborators who have high number of papers and high citations per paper tends to promote the positive growth of a researcher's relative role at that stage, thereby promoting the formation of positive growth category and vice versa, which is consistent with previous studies demonstrating that co-authorship with influential collaborators has a positive impact on career success (Li et al., 2019; Liénard et al., 2018; Malmgren et al., 2010; Peterson, 2015).

To sum up, this article has some research and policy implications. We investigate the relative role growth patterns and the influence of collaborators on the formation of growth patterns, which can help understand the essential characteristics of academic career development and provide guidance on academic strategies in choosing collaborators. Besides, the findings of research performance under diverse growth types in this work can actually assist researchers in formulating more favorable career development plans and reasonable collaborator selection strategy during their careers; meanwhile, it provides ideas for governments and universities to formulate scientific research policies and cultivate principal investigators.

There are, however, some limitations in this work. First of all, the amount of data we selected is not large enough, as it only includes 1315 authors; we will expand the conditions to allow more scholars to be included in our further research. What's more, the relative role growth patterns of the academic career lack comparison between different fields, so the research may not be universally applicative. Last but not least, only the number of coauthors participates in our work when analyzing the relative role of the researchers, therefore, we will explore institutions of authors in our future work.

Other related research can be further explored in the future. First, we use the h-index indicator to calculate the classification indicator of relative role; future research can consider using multiple indicators for calculation and comprehensive comparative analysis. Second, we analyze relative role growth patterns and divide these patterns into four growth types based on career development characteristics; future research can classify and explore career growth patterns from another dimension. Then, the regressive growth type of collaboration showing less productivity includes a possibility that some senior scientists may collaborate with more senior authors in their late career, which deserves further investigation of this phenomenon. Finally, we investigate some indicators of collaborative ability and research performance in diverse relative role growth types; future work can compare and analyze more evaluating indicators. In the future, we will continue to conduct analysis of collaboration patterns and motives by adding qualitative studies such as a survey.

Author contributions

Wei Lu: Conceived and designed the analysis, Contributed data or analysis tools, Performed the analysis.
Yan Ren: Conceived and designed the analysis, Contributed data or analysis tools, Performed the analysis, Wrote the paper.
Yong Huang: Conceived and designed the analysis, Contributed data or analysis tools, Performed the analysis, Wrote the paper.
Yi Bu: Conceived and designed the analysis, Wrote the paper.
Yuehan Zhang: Collected the data, Performed the analysis.

Acknowledgements

This work was supported by the Youth Science Foundation of the National Natural Science Foundation of China (grant no. 72004168).The authors are grateful to the anonymous referees and editors for their invaluable and insightful comments.

References

Abbasi, A., Chung, K. S. K., & Hossain, L. (2012). Egocentric analysis of co-authorship network structure, position and performance. *Information Processing & Management, 48*(4), 671–679.
Badar, K., Frantz, T. L., & Jabeen, M. (2016). Research performance and degree centrality in co-authorship networks. *Aslib Journal of Information Management*.
Bai, X., Zhang, F., & Lee, I. (2019). Predicting the citations of scholarly paper. *Journal of Informetrics, 13*(1), 407–418.
Beaver, D. B., & Rosen, R. (1978). Studies in scientific collaboration: part I. The professional origins of scientific co-authorship. *Scientometrics, 1*(3), 65–84.
Bol, T., de Vaan, M., & van de Rijt, A. (2018). The Matthew effect in science funding. *Proceedings of the National Academy of Sciences, 115*(19), 4887–4890.
Bordons, M., Aparicio, J., González-Albo, B., & Díaz-Faes, A. A. (2015). The relationship between the research performance of scientists and their position in co-authorship networks in three fields. *Journal of informetrics, 9*(1), 135–144.
Bozeman, B., & Corley, E. (2004). Scientists' collaboration strategies: implications for scientific and technical human capital. *Research Policy, 33*(4), 599–616.
Cronin, B., & Meho, L. (2006). Using the h-index to rank influential information scientistss. *Journal of the American Society for Information Science and Technology, 57*(9), 1275–1278.
De Solla Price, D. J. (1963). *Little science, big science*. New York: Columbia University Press.
Drożdż, S., Kulig, A., Kwapień, J., Niewiarowski, A., & Stanuszek, M. (2017). Hierarchical organization of H. Eugene Stanley scientific collaboration community in weighted network representation. *Journal of Informetrics, 11*(4), 1114–1127.
Everett, M., & Borgatti, S. P. (2005). Ego-centric network betweenness. *Social Networks, 27*(1), 31–38.
Feeney, M. K., & Welch, E. W. (2014). Academic outcomes among principal investigators, co-principal investigators, and non-PI researchers. *The Journal of Technology Transfer, 39*(1), 111–133.
Glänzel, W. (2014). Analysis of co-authorship patterns at the individual level. *Transinformação, 26*(3), 229–238.
He, Z.-L., Geng, X.-S., & Campbell-Hunt, C. (2009). Research collaboration and research output: A longitudinal study of 65 biomedical scientists in a New Zealand university. *Research Policy, 38*(2), 306–317.
Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences, 102*(46), 16569–16572.
Kastrin, A., Klisara, J., Lužar, B., & Povh, J. (2018). Is science driven by principal investigators? *Scientometrics, 117*(2), 1157–1182.
Kong, X., Mao, M., Jiang, H., Yu, S., & Wan, L. (2019). How does collaboration affect researchers' positions in co-authorship networks? *Journal of Informetrics, 13*(3), 887–900.
Kong, X., Zhang, J., Zhang, D., Bu, Y., Ding, Y., & Xia, F. (2020). The Gene of Scientific Success. *ACM Transactions on Knowledge Discovery from Data (TKDD), 14*(4), 1–19.
Lande, D., Fu, M., Guo, W., Balagura, I., Gorbov, I., & Yang, H. (2020). Link prediction of scientific collaboration networks based on information retrieval. *World Wide Web*, 1–19.

Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science, 35*(5), 673–702.
Liénard, Jean F, Achakulvisut, T., Acuna, D. E., & David, S. V. (2018). Intellectual synthesis in mentorship determines success in academic careers. *Nature Communications, 9*(1).
Li, N., & Gillet, D. (2013). Identifying influential scholars in academic social media platforms. In *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (pp. 608–614).
Li, W., Aste, T., Caccioli, F., & Livan, G. (2019). Beginning co-authorship with top scientists predicts success in academic careers. *Nature Communications, 10*(1), 1–9.
Li, Y., Li, H., Liu, N., & Liu, X. (2018). Important institutions of interinstitutional scientific collaboration networks in materials science. *Scientometrics, 117*(1), 85–103.
Malmgren, R. D., Ottino, J. M., & Amaral, L. A. N. (2010). The role of mentorship in protégé performance. *Nature, 465*(7298), 622–626.
Melkers, J., & Xiao, F. (2012). Boundary-spanning in emerging technology research: Determinants of funding success for academic scientists. *The Journal of Technology Transfer, 37*(3), 251–270.
Merton, R. K. (1957). Priorities in scientific discovery: a chapter in the sociology of science. *American Sociological Review, 22*(6), 635–659.
Milojević, S., Radicchi, F., & Walsh, J. P. (2018). Changing demographics of scientific careers: The rise of the temporary workforce. *Proceedings of the National Academy of Sciences, 115*(50), 12616–12623.
Ortega, J. L. (2014). Influence of co-authorship networks in the research impact: Ego-centric network analyses from Microsoft Academic Search. *Journal of Informetrics, 8*(3), 728–737.
Petersen, A. M. (2015). Quantifying the impact of weak, strong, and super ties in scientific careers. *Proceedings of the National Academy of Sciences, 112*(34), E4671–E4680.
Sugimoto, C. R., Sugimoto, T. J., Tsou, A., Milojević, S., & Larivière, V. (2016). Age stratification and cohort effects in scholarly communication: a study of social sciences. *Scientometrics, 109*(2), 997–1016.
Tol, R. S. (2013). The Matthew effect for cohorts of economists. *Journal of Informetrics, 7*(2), 522–527.
Wahba, M. A., & Bridwell, L. (1973). Maslow's need hierarchy theory: A review of research. In *Proceedings of the Annual Convention of the* (pp. 571–572). American Psychological Association.
Wang, W., Yu, S., Bekele, T. M., Kong, X., & Xia, F. (2017). Scientific collaboration patterns vary with scholars' academic ages. *Scientometrics, 112*(1), 329–343.
Wang, Y., Jones, B. F., & Wang, D. (2019). Beginning-career setback and future career impact. *Nature Communications, 10*(1), 1–10.
Yin, Z., & Zhi, Q. (2017). Dancing with the academic elite: a promotion or hindrance of research production? *Scientometrics, 110*(1), 17–41.
Yoshikane, F., Nozawa, T., & Tsuji, K. (2006). Comparative analysis of co-authorship networks considering authors' roles in collaboration: differences between the theoretical and application areas. *Scientometrics, 68*(3), 643–655.
Yoshikane, F., Nozawa, T., Shibui, S., & Suzuki, T. (2009). An analysis of the connection between researchers' productivity and their co-authors' past attributions, including the importance in co-authorship networks. *Scientometrics, 79*(2), 435–449.