Trabalho final - NT265

Mariane Font Fernandes

17/05/2019

## Introdução

O desenvolvimento de um projeto de pesquisa na pós-graduação requer do estudante a leitura constante de artigos que ampliem seu conhecimento específico, sendo necessário saber a tendência da comunidade científica e o quanto já foi discutido sobre o assunto. A busca por autores bem conceituados na área, assim como artigos de grande impacto para citação, não é fácil em meio a tantos materiais disponíveis. Um modo de conseguir um direcionamento seria através do R e seu pacote “Bibliometrix”, que é capaz de fazer análises quantitativas e estatísticas de publicações e suas citações. Para exemplificar, este roteiro mostra a análise de artigos, que possuem crotonilação (*crotonylation* em inglês) em seu título, seguindo os códigos disponíveis neste pacote.

## Requerimentos:

Para que este roteiro seja executado, é necessário fazer download do:

* [Software R](https://cran.r-project.org/bin/windows/base/);
* [RStudio](https://www.rstudio.com/products/rstudio/download/);
* [Git](https://git-scm.com/downloads);
* [MikTex](https://miktex.org/download);
* Pacote [Rmarkdown](https://cran.r-project.org/web/packages/rmarkdown/index.html), escrevendo no console do RStudio install.packages(“rmarkdown”);
* Pacote [Bibliometrix](https://cran.r-project.org/web/packages/bibliometrix/index.html), escrevendo no console do RStudio install.packages(“bibliometrix”);
* Arquivo Scopus.bib, com dados dos artigos relacionados a crotonilação, adquirido a partir da base de dados [Scopus](https://www.scopus.com/).

## Desenvolvimento:

### 1. Carregando dados e covertendo

Após fazer o download das ferramentas necessárias e criar um diretório no GitHub, carregamos o pacote necessário para análise no RStudio com o comando:

library(bibliometrix)

## Registered S3 methods overwritten by 'ggplot2':  
## method from   
## [.quosures rlang  
## c.quosures rlang  
## print.quosures rlang

## To cite bibliometrix in publications, please use:  
##   
## Aria, M. & Cuccurullo, C. (2017) bibliometrix: An R-tool for comprehensive science mapping analysis, Journal of Informetrics, 11(4), pp 959-975, Elsevier.  
##   
##   
## http:\\www.bibliometrix.org  
##   
##   
## To start with the shiny web-interface, please digit:  
## biblioshiny()

E carregamos o arquivo Scopus.bib, que será atribuído ao vetor D através da função “readFiles”:

D <- readFiles("Scopus.bib")

Essa função é capaz de juntar outros arquivos de texto em um mesmo vetor e os converte em UTF-8. Como precisamos dele em data frame para realizar a análise, usaremos a função “convert2df”, indicando o banco de dados utilizado em “dbsource” e o formato do arquivo em “format”. Armazenaremos os dados convertivos em M:

M <- convert2df(D, dbsource="scopus", format="bibtex")

##   
## Converting your scopus collection into a bibliographic dataframe  
##   
## Articles extracted 36   
## Done!  
##   
##   
## Generating affiliation field tag AU\_UN from C1: Done!

Em caso de dúvidas quanto as funções, é possível buscar ajuda no RStudio através do comando ?nome\_da\_função

### 2. Análise bibliométrica

Para fazer a análise descritiva dos dados, utilizaremos a função “biblioAnalysis”, que retorna as principais medidas na bibliometria em um objeto da classe “bibliometrix”:

results <- biblioAnalysis(M, sep=";")

### 3. Função *summary* e *plot*

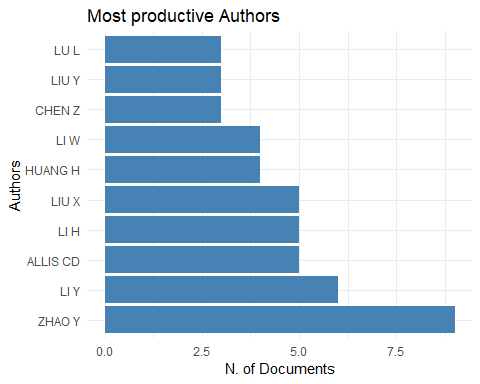
Para resumir os resultados bibliométricos obtidos e conseguir suas principais informações, assim como produção científica anual, principais artigos por número de citações, autores e países mais produtivos, citação total por país, fontes e palavras-chave mais relevantes, usamos a função “summary”:

options(width=100)  
S <- summary(object=results, k=10, pause=FALSE)

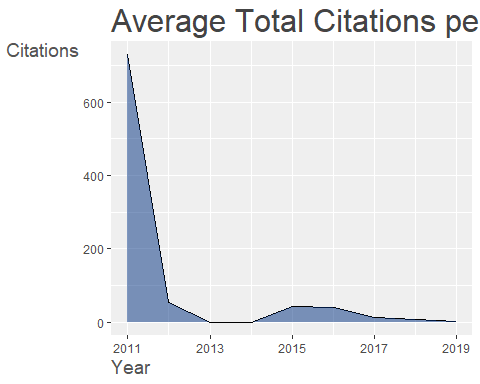
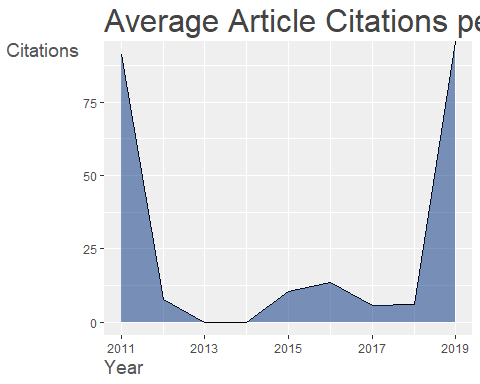
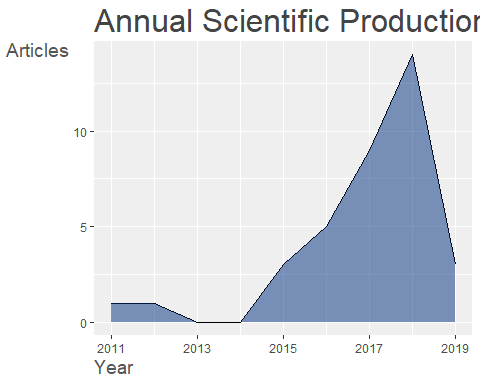
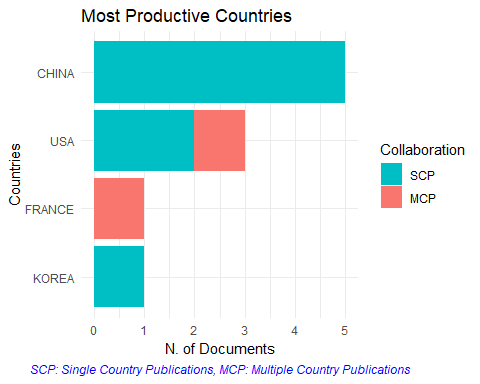
##   
##   
## Main Information about data  
##   
## Documents 36   
## Sources (Journals, Books, etc.) 24   
## Keywords Plus (ID) 590   
## Author's Keywords (DE) 58   
## Period 2011 - 2019   
## Average citations per documents 36.19   
##   
## Authors 226   
## Author Appearances 311   
## Authors of single-authored documents 1   
## Authors of multi-authored documents 225   
## Single-authored documents 1   
##   
## Documents per Author 0.159   
## Authors per Document 6.28   
## Co-Authors per Documents 8.64   
## Collaboration Index 6.43   
##   
## Document types   
## ARTICLE 30   
## ERRATUM 1   
## LETTER 2   
## NOTE 2   
## REVIEW 1   
##   
##   
## Annual Scientific Production  
##   
## Year Articles  
## 2011 1  
## 2012 1  
## 2015 3  
## 2016 5  
## 2017 9  
## 2018 14  
## 2019 3  
##   
## Annual Percentage Growth Rate 20.0937   
##   
##   
## Most Productive Authors  
##   
## Authors Articles Authors Articles Fractionalized  
## 1 ZHAO Y 9 ZHAO Y 1.211  
## 2 LI Y 6 BAUMANN K 1.000  
## 3 ALLIS CD 5 LI Y 0.755  
## 4 LI H 5 LI H 0.655  
## 5 LIU X 5 HUANG H 0.571  
## 6 HUANG H 4 BOS J 0.500  
## 7 LI W 4 CHAN JC 0.500  
## 8 CHEN Z 3 HE JJ 0.500  
## 9 LIU Y 3 JU Z 0.500  
## 10 LU L 3 MAZE I 0.500  
##   
##   
## Top manuscripts per citations  
##   
## Paper TC TCperYear  
## 1 TAN M, 2011, CELL 729 91.12  
## 2 SABARI BR, 2015, MOL CELL-a 121 30.25  
## 3 LI Y, 2016, MOL CELL 71 23.67  
## 4 MONTELLIER E, 2012, BIOESSAYS 54 7.71  
## 5 ANDREWS FH, 2016, NAT CHEM BIOL 49 16.33  
## 6 QIU WR, 2018, GENOMICS 42 42.00  
## 7 ZHAO D, 2016, CELL RES 33 11.00  
## 8 XIONG X, 2016, NAT CHEM BIOL 32 10.67  
## 9 FELLOWS R, 2018, NAT COMMUN 26 26.00  
## 10 WEI W, 2017, CELL RES 21 10.50  
##   
##   
## Corresponding Author's Countries  
##   
## Country Articles Freq SCP MCP MCP\_Ratio  
## 1 CHINA 5 0.5 5 0 0.000  
## 2 USA 3 0.3 2 1 0.333  
## 3 FRANCE 1 0.1 0 1 1.000  
## 4 KOREA 1 0.1 1 0 0.000  
##   
##   
## SCP: Single Country Publications  
##   
## MCP: Multiple Country Publications  
##   
##   
## Total Citations per Country  
##   
## Country Total Citations Average Article Citations  
## 1 USA 735 245  
## 2 FRANCE 54 54  
## 3 CHINA 45 9  
## 4 KOREA 1 1  
##   
##   
## Most Relevant Sources  
##   
## Sources Articles  
## 1 MOLECULAR CELL 4  
## 2 SCIENTIFIC REPORTS 4  
## 3 CELL RESEARCH 3  
## 4 JOURNAL OF PROTEOME RESEARCH 3  
## 5 BIOLOGICAL PSYCHIATRY 2  
## 6 NATURE CHEMICAL BIOLOGY 2  
## 7 ARTIFICIAL INTELLIGENCE IN MEDICINE 1  
## 8 BIOESSAYS 1  
## 9 BIOMEDICINE AND PHARMACOTHERAPY 1  
## 10 CELL 1  
##   
##   
## Most Relevant Keywords  
##   
## Author Keywords (DE) Articles Keywords-Plus (ID) Articles  
## 1 LYSINE CROTONYLATION 6 LYSINE 44  
## 2 CROTONYLATION 3 PROTEIN PROCESSING 36  
## 3 HISTONE CROTONYLATION 3 HISTONE 34  
## 4 POST TRANSLATIONAL MODIFICATION 3 HUMAN 31  
## 5 ACETYLATION 2 ARTICLE 22  
## 6 CDYL 2 HISTONES 20  
## 7 MASS SPECTROMETRY 2 METABOLISM 20  
## 8 PTMS 2 NONHUMAN 17  
## 9 SAHA 2 POST TRANSLATIONAL 16  
## 10 SUPPORT VECTOR MACHINE 2 PRIORITY JOURNAL 16

Para ver alguns gráficos básicos, usamos a função “plot”, que assim como summary, aceita dois argumentos, sendo o “k” o número de linhas de cada tabela e “pause” um valor lógico (TRUE ou FALSE) que define se haverá rolagem de tela ou não.

plot(x=results, k=10, pause=FALSE)



## Warning: Removed 12 rows containing missing values (position\_stack).



### 4. Análise de referências citadas

Para ter acesso a uma tabela com a frequência das referências ou principais autores mais citados, utilizamos a função “citations”, sendo necessário indicar qual das duas informações queremos em “field” e qual o separador utilizado em “sep”. Para confirmar qual o separador, utilizamos o seguinte código:

M$CR[1]

## [1] "SUN, H., KENNEDY, P.J., NESTLER, E.J., EPIGENETICS OF THE DEPRESSED BRAIN: ROLE OF HISTONE ACETYLATION AND METHYLATION (2013) NEUROPSYCHOPHARMACOLOGY, 38, PP. 124-137; LIU, Y., LI, M., FAN, M., SONG, Y., YU, H., ZHI, X., CHROMODOMAIN Y-LIKE PROTEINMEDIATED HISTONE CROTONYLATION REGULATES STRESS-INDUCED DEPRESSIVE BEHAVIORS (2019) BIOL PSYCHIATRY, 85, PP. 635-649; TAN, M., LUO, H., LEE, S., JIN, F., YANG, J.S., MONTELLIER, E., IDENTIFICATION OF 67 HISTONE MARKS AND HISTONE LYSINE CROTONYLATION AS A NEW TYPE OF HISTONE MODIFICATION (2011) CELL, 146, PP. 1016-1028; SABARI, B.R., ZHANG, D., ALLIS, C.D., ZHAO, Y., METABOLIC REGULATION OF GENE EXPRESSION THROUGH HISTONE ACYLATIONS (2017) NAT REV MOL CELL BIOL, 18, PP. 90-101; FELLOWS, R., DENIZOT, J., STELLATO, C., CUOMO, A., JAIN, P., STOYANOVA, E., MICROBIOTA DERIVED SHORT CHAIN FATTY ACIDS PROMOTE HISTONE CROTONYLATION IN THE COLON THROUGH HISTONE DEACETYLASES (2018) NAT COMMUN, 9, P. 105; SABARI, B.R., TANG, Z., HUANG, H., YONG-GONZALEZ, V., MOLINA, H., KONG, H.E., INTRACELLULAR CROTONYL-COA STIMULATES TRANSCRIPTION THROUGH P300-CATALYZED HISTONE CROTONYLATION (2015) MOL CELL, 58, PP. 203-215; LI, Y., SABARI, B.R., PANCHENKO, T., WEN, H., ZHAO, D., GUAN, H., MOLECULAR COUPLING OF HISTONE CROTONYLATION AND ACTIVE TRANSCRIPTION BY AF9 YEATS DOMAIN (2016) MOL CELL, 62, PP. 181-193; SIMITHY, J., SIDOLI, S., YUAN, Z.F., CORADIN, M., BHANU, N.V., MARCHIONE, D.M., CHARACTERIZATION OF HISTONE ACYLATIONS LINKS CHROMATIN MODIFICATIONS WITH METABOLISM (2017) NAT COMMUN, 8, P. 1141; LIU, S., YU, H., LIU, Y., LIU, X., ZHANG, Y., BU, C., CHROMODOMAIN PROTEIN CDYL ACTS AS A CROTONYL-COA HYDRATASE TO REGULATE HISTONE CROTONYLATION AND SPERMATOGENESIS (2017) MOL CELL, 67, PP. 853-866; NETO, F.L., BORGES, G., TORRES-SANCHEZ, S., MICO, J.A., BERROCOSO, E., NEUROTROPHINS ROLE IN DEPRESSION NEUROBIOLOGY: A REVIEW OF BASIC AND CLINICAL EVIDENCE (2011) CURR NEUROPHARMACOL, 9, PP. 530-552"

A figura mostra as referências do primeiro artigo. Nesse caso, o campo separador é “;”.

Para obter os artigos mais citados então, usamos o seguinte código:

CR <- citations(M, field="article", sep=";")  
cbind(CR$Cited[1:10])

## [,1]  
## TAN, M., LUO, H., LEE, S., JIN, F., YANG, J.S., MONTELLIER, E., BUCHOU, T., RAJAGOPAL, N., IDENTIFICATION OF 67 HISTONE MARKS AND HISTONE LYSINE CROTONYLATION AS A NEW TYPE OF HISTONE MODIFICATION (2011) CELL, 146, PP. 1016-1028 6  
## TAN, M., IDENTIFICATION OF 67 HISTONE MARKS AND HISTONE LYSINE CROTONYLATION AS A NEW TYPE OF HISTONE MODIFICATION (2011) CELL, 146, PP. 1016-1028 5  
## BERDASCO, M., ESTELLER, M., ABERRANT EPIGENETIC LANDSCAPE IN CANCER: HOW CELLULAR IDENTITY GOES AWRY (2010) DEV. CELL, 19, PP. 698-711 4  
## BAO, X., WANG, Y., LI, X., LI, X.M., LIU, Z., YANG, T., WONG, C.F., LI, X.D., IDENTIFICATION OF 'ERASERS' FOR LYSINE CROTONYLATED HISTONE MARKS USING A CHEMICAL PROTEOMICS APPROACH (2014) ELIFE, 3, P. E02999 3  
## KOUZARIDES, T., CHROMATIN MODIFICATIONS AND THEIR FUNCTION (2007) CELL, 128, PP. 693-705 3  
## LI, Y., SABARI, B.R., PANCHENKO, T., WEN, H., ZHAO, D., GUAN, H., MOLECULAR COUPLING OF HISTONE CROTONYLATION AND ACTIVE TRANSCRIPTION BY AF9 YEATS DOMAIN (2016) MOL CELL, 62, PP. 181-193 3  
## MONTELLIER, E., ROUSSEAUX, S., ZHAO, Y., KHOCHBIN, S., HISTONE CROTONYLATION SPECIFICALLY MARKS THE HAPLOID MALE GERM CELL GENE EXPRESSION PROGRAM: POST-MEIOTIC MALE-SPECIFIC GENE EXPRESSION (2012) BIOESSAYS, 34, PP. 187-193 3  
## SABARI, B.R., TANG, Z., HUANG, H., YONG-GONZALEZ, V., MOLINA, H., KONG, H.E., DAI, L., ZHAO, Y., INTRACELLULAR CROTONYL-COA STIMULATES TRANSCRIPTION THROUGH P300-CATALYZED HISTONE CROTONYLATION (2015) MOL. CELL, 58, PP. 203-215 3  
## TAN, M., LUO, H., LEE, S., JIN, F., YANG, J.S., MONTELLIER, E., IDENTIFICATION OF 67 HISTONE MARKS AND HISTONE LYSINE CROTONYLATION AS A NEW TYPE OF HISTONE MODIFICATION (2011) CELL, 146, PP. 1016-1028 3  
## TAVERNA, S.D., LI, H., RUTHENBURG, A.J., ALLIS, C.D., PATEL, D.J., HOW CHROMATIN-BINDING MODULES INTERPRET HISTONE MODIFICATIONS: LESSONS FROM PROFESSIONAL POCKET PICKERS (2007) NAT. STRUCT. MOL. BIOL., 14, PP. 1025-1040 3

Para obter os autores mais citados:

CR <- citations(M, field="author", sep=";")  
cbind(CR$Cited[1:10])

## [,1]  
## ZHAO Y 90  
## SABARI B R 54  
## CHOU K C 51  
## CHEN Y 50  
## TAN M 49  
## ALLIS C D 45  
## XIAO X 44  
## MONTELLIER E 42  
## HUANG H 39  
## DAI L 37

Para saber quantas vezes um autor (ou um documento) em uma coleção foi citado por outros autores na mesma coleção, utilizamos a função “localCitations”:

CR <- localCitations(M, sep=";")

## Articles analysed 29

CR$Authors[1:10,]

## Author LocalCitations  
## 197 ZHAO Y 83  
## 1 ALLIS CD 49  
## 53 HUANG H 37  
## 113 ROEDER RG 36  
## 117 SABARI BR 36  
## 143 TANG Z 36  
## 81 LI Y 34  
## 67 KHOCHBIN S 33  
## 101 MONTELLIER E 33  
## 114 ROUSSEAUX S 33

CR$Papers[1:10,]

## Paper DOI Year LCS GCS  
## 1 TAN M, 2011, CELL 10.1016/J.CELL.2011.08.008 2011 28 729  
## 3 SABARI BR, 2015, MOL CELL-a 10.1016/J.MOLCEL.2015.02.029 2015 21 121  
## 9 LI Y, 2016, MOL CELL 10.1016/J.MOLCEL.2016.03.028 2016 15 71  
## 16 WEI W, 2017, CELL RES 10.1038/CR.2017.68 2017 10 21  
## 8 ZHAO D, 2016, CELL RES 10.1038/CR.2016.49 2016 8 33  
## 17 WEI W, 2017, J PROTEOME RES 10.1021/ACS.JPROTEOME.7B00012 2017 8 16  
## 6 RUIZ-ANDRES O, 2016, DNM DIS MODELS MECH 10.1242/DMM.024455 2016 7 20  
## 7 ANDREWS FH, 2016, NAT CHEM BIOL 10.1038/NCHEMBIO.2065 2016 7 49  
## 15 XU W, 2017, CELL RES 10.1038/CR.2017.60 2017 7 9  
## 2 MONTELLIER E, 2012, BIOESSAYS 10.1002/BIES.201100141 2012 5 54

### 5. Ranking de dominância de autores

O ranking de dominância dos autores, feito pela função “dominance”, indica a fração de artigos em que um determinado pesquisador aparece como primeiro autor. Essa função depende do objeto obtido pela biblioAnalysis e “k” indicando o número de autores a considerar na análise.

DF <- dominance(results, k = 10)  
DF

## Author Dominance Factor Tot Articles Single-Authored Multi-Authored First-Authored Rank by Articles Rank by DF  
## 1 LIU S 1.0000000 2 0 2 2 1 1  
## 2 QIU WR 1.0000000 2 0 2 2 1 1  
## 3 WEI W 1.0000000 2 0 2 2 1 1  
## 4 SABARI BR 0.6666667 3 0 3 2 6 4  
## 5 MONTELLIER E 0.5000000 2 0 2 1 1 5  
## 6 WAN J 0.5000000 2 0 2 1 1 5  
## 7 LI Y 0.3333333 6 0 6 2 10 7  
## 8 LIU Y 0.3333333 3 0 3 1 6 7  
## 9 ZHAO D 0.3333333 3 0 3 1 6 7  
## 10 HUANG H 0.2500000 4 0 4 1 9 10

Neste exemplo, LIU, QIU e WEI dominam seus grupos de pesquisa, por só terem artigos como primeiro autor.

### 6. Índice-h de autores

Para medir o impacto da produtividade e da citação das publicações de um pesquisador, utilizamos a função “Hindex”, calculando o índice H dos autores ou das fontes e suas variantes (índice-g e índice-m) em uma coleção bibliográfica. Essa função tem como argumento um dado bibliográfico, o elemento que queremos analisar em “field” e nomes em “elements”, separando o sobrenome e as iniciais por um espaço em branco para autores.

indices <- Hindex(M, field = "author", elements="LI Y", sep = ";")  
  
# índice de impacto de LI Y:  
indices$H

## Author h\_index g\_index m\_index TC NP PY\_start  
## 1 LI Y 5 6 1.25 151 6 2016

# Citações de LI Y:  
indices$CitationList

## [[1]]  
## Authors Journal Year TotalCitation  
## 1 YANG Q;LI Y;APALIYA MT;ZHENG X FRONTIERS IN MICROBIOLOGY 2018 0  
## 3 LI Y;ZHAO D;CHEN Z;LI H TRANSCRIPTION 2017 7  
## 2 SUN H;LIU X;LI F;LI W;ZHANG J; SCIENTIFIC REPORTS 2017 8  
## 4 XIONG X;PANCHENKO T;YANG S;ZHA NATURE CHEMICAL BIOLOGY 2016 32  
## 5 ZHAO D;GUAN H;ZHAO S;MI W;WEN CELL RESEARCH 2016 33  
## 6 LI Y;SABARI BR;PANCHENKO T;WEN MOLECULAR CELL 2016 71

Para calcular o índice h dos primeiros 10 autores mais produtivos (nesta coleção) usamos o seguinte código:

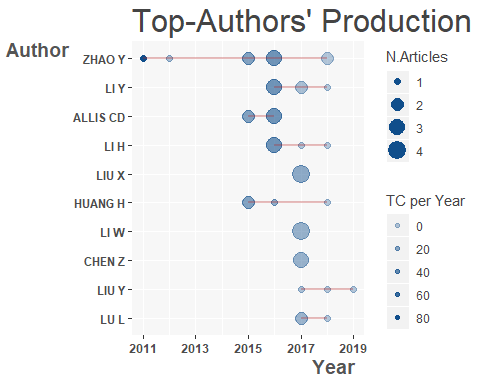
authors=gsub(","," ",names(results$Authors)[1:10])  
  
indices <- Hindex(M, field = "author", elements=authors, sep = ";", years = 50)  
  
indices$H

## Author h\_index g\_index m\_index TC NP PY\_start  
## 1 ZHAO Y 6 9 0.6666667 1041 9 2011  
## 2 LI Y 5 6 1.2500000 151 6 2016  
## 3 ALLIS CD 4 5 0.8000000 257 5 2015  
## 4 LI H 4 5 1.0000000 144 5 2016  
## 5 LIU X 4 4 1.3333333 58 4 2017  
## 6 HUANG H 2 4 0.4000000 193 4 2015  
## 7 LI W 4 4 1.3333333 44 4 2017  
## 8 CHEN Z 3 3 1.0000000 41 3 2017  
## 9 LIU Y 1 3 0.3333333 14 3 2017  
## 10 LU L 2 3 0.6666667 38 3 2017

### 7. Produtividade dos principais autores ao decorrer do tempo

A função “AuthorProdOverTime” calcula e plota a produção dos autores ao longo do tempo, de acordo com o número de publicações e o total de citações por ano. Os argumentos necessários são os dados bibliográficos, o número de autores que você deseja ver em “k” e um valor lógico (TRUE ou FALSE) em “graph” para plotar ou não um gráfico de produção do autor ao longo do tempo.

topAU <- authorProdOverTime(M, k=10, graph=TRUE)



Para saber a produtividade dos autores por ano:

head(topAU$dfAU)

## Author year freq TC TCpY  
## 1 ALLIS CD 2015 2 121 24.20000  
## 2 ALLIS CD 2016 3 136 34.00000  
## 3 CHEN Z 2017 3 41 13.66667  
## 4 HUANG H 2015 2 121 24.20000  
## 5 HUANG H 2016 1 71 17.75000  
## 6 HUANG H 2018 1 1 0.50000

Para saber a lista de artigos desses autores:

head(topAU$dfPapersAU)

## Author year  
## 2 ZHAO Y 2018  
## 3 ZHAO Y 2018  
## 4 ZHAO Y 2016  
## 5 ZHAO Y 2016  
## 6 ZHAO Y 2016  
## 7 ZHAO Y 2015  
## TI  
## 2 DYNAMICS AND FUNCTIONAL INTERPLAY OF HISTONE LYSINE BUTYRYLATION, CROTONYLATION, AND ACETYLATION IN RICE UNDER STARVATION AND SUBMERGENCE  
## 3 QUANTITATIVE CROTONYLOME ANALYSIS EXPANDS THE ROLES OF P300 IN THE REGULATION OF LYSINE CROTONYLATION PATHWAY  
## 4 SELECTIVE RECOGNITION OF HISTONE CROTONYLATION BY DOUBLE PHD FINGERS OF MOZ AND DPF2  
## 5 YEATS2 IS A SELECTIVE HISTONE CROTONYLATION READER  
## 6 MOLECULAR COUPLING OF HISTONE CROTONYLATION AND ACTIVE TRANSCRIPTION BY AF9 YEATS DOMAIN  
## 7 ERRATUM: INTRACELLULAR CROTONYL-COA STIMULATES TRANSCRIPTION THROUGH P300-CATALYZED HISTONE CROTONYLATION (MOLECULAR CELL (2015) 58(2) (203215)(S1097276515001434)(10.1016/J.MOLCEL.2015.02.029))  
## SO DOI TC TCpY  
## 2 GENOME BIOLOGY 10.1186/S13059-018-1533-Y 0 0.00  
## 3 PROTEOMICS 10.1002/PMIC.201700230 1 0.50  
## 4 NATURE CHEMICAL BIOLOGY 10.1038/NCHEMBIO.2218 32 8.00  
## 5 CELL RESEARCH 10.1038/CR.2016.49 33 8.25  
## 6 MOLECULAR CELL 10.1016/J.MOLCEL.2016.03.028 71 17.75  
## 7 MOLECULAR CELL 10.1016/J.MOLCEL.2018.01.013 0 0.00

### 8. Estimativa de coeficiente pela lei de Lotka

A função “lotka” verifica a produtividade científica através da lei de Lotka, que descreve a frequência de publicação dos autores como uma lei do inverso do quadrado, onde a quantidade de autores publicando um certo número de artigos é uma razão fixa para a quantidade de autores publicando um único artigo, sendo o coeficiente beta da lei de Lotka igual a 2. É possível avaliar a similaridade desta distribuição empírica, dada pelo coeficiente beta, com a distribuição teórica através de um teste estatístico.

L <- lotka(results)

Para saber a distribuição empirica da produtividade do autor:

L$AuthorProd

## N.Articles N.Authors Freq  
## 1 1 174 0.769911504  
## 2 2 36 0.159292035  
## 3 3 9 0.039823009  
## 4 4 2 0.008849558  
## 5 5 3 0.013274336  
## 6 6 1 0.004424779  
## 7 9 1 0.004424779

Coeficiente Beta estimado:

L$Beta

## [1] 2.577603

Constante:

L$C

## [1] 0.6907428

Qualidade do ajuste:

L$R2

## [1] 0.9416943

p-valor obtido no teste de Kolmogorov-Smirnov:

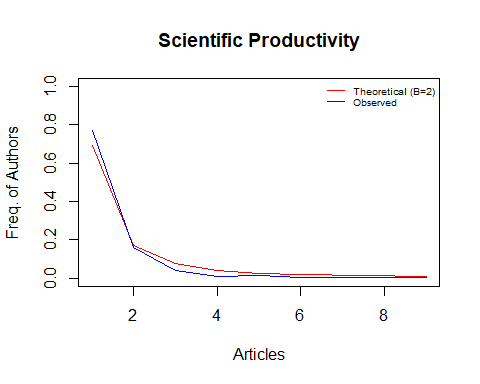
L$p.value

## [1] 0.5412431

O teste de Kolmogorov-Smirnov fornece um valor de p de 0,54, o que significa que não há diferença significativa entre as distribuições de Lotka observadas e as teóricas.

Podemos comprovar esse resultado comparando as duas distribuições pela função de plotagem:

# Distribuição observada:  
Observed=L$AuthorProd[,3]  
  
# Distribuição teórica com Beta = 2  
Theoretical=10^(log10(L$C)-2\*log10(L$AuthorProd[,1]))  
  
plot(L$AuthorProd[,1],Theoretical,type="l",col="red",ylim=c(0, 1), xlab="Articles",ylab="Freq. of Authors",main="Scientific Productivity")  
lines(L$AuthorProd[,1],Observed,col="blue")  
legend(x="topright",c("Theoretical (B=2)","Observed"),col=c("red","blue"),lty = c(1,1,1),cex=0.6,bty="n")



### 9. Matrizes de redes bibliográficas

Os atributos dos artigos estão conectados entre si, gerando redes bipartidas que podem ser representadas como matrizes retangulares (Artigos x Atributos).

Além disso, as referências feitas em publicações científicas gera uma rede adicional, chamada de rede de co-citação ou acoplamento.

Essas redes são analisadas para determinar propriedades significativas do sistema de pesquisa subjacente e verificar a influência de unidades bibliométricas, como acadêmicos e periódicos.

#### a. Redes bipartidas

Para calcular uma rede bipartida selecionando um de seus atributos, utilizamos a função “cocMatrix”.

U <- cocMatrix(M, Field = "SO", sep = ";")

Neste caso, relacionamos artigos (linhas) e fontes de publicação (colunas) em uma rede bipartida, usando “SO” em “fields”.

Classificando, em ordem decrescente, podemos ver as fontes de publicação mais relevantes:

sort(Matrix::colSums(U), decreasing = TRUE)[1:5]

## SCIENTIFIC REPORTS MOLECULAR CELL JOURNAL OF PROTEOME RESEARCH CELL RESEARCH   
## 4 4 3 3   
## BIOLOGICAL PSYCHIATRY   
## 2

Outras possibilidades para o campo “fields” incluem “CR” para rede de citações, “AU” para rede de autores, “DE” para rede de palavras-chave do autor, “ID” para rede de mais palavras-chave e “AU\_CO” para rede de países dos autores. Neste último caso, precisamos extrair essas informações do atributo de afiliação usando a função “metaTagExtraction”, por não ser um atributo padrão de tabelas de dados bibliográficos.

# Rede de Citações:  
X <- cocMatrix(M, Field="CR", sep=". ")

# Rede de Autores:  
Y <- cocMatrix(M, Field = "AU", sep = ";")

# Rede de Palavras-chave do Autor:  
V <- cocMatrix(M, Field = "DE", sep = ";")  
V

## 36 x 58 sparse Matrix of class "dgCMatrix"

## [[ suppressing 58 column names 'CDYL', 'DENDRITIC SPINE', 'DEPRESSION' ... ]]

##   
## CHAN JC, 2019, BIOL PSYCHIATRY . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## LIU Y, 2019, BIOL PSYCHIATRY 1 1 1 1 1 1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## WAN J, 2019, BIOMED PHARMACOTHER . . . . . . 1 1 1 . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## KELLY RDW, 2018, SCI REP . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## LIU K, 2018, SCI REP . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## KWON OK, 2018, SCI REP . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## FELLOWS R, 2018, NAT COMMUN . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## LIU S, 2018, MOL CELL PROTEOMICS . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## LU Y, 2018, GENOME BIOL . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## WU Q, 2018, J PROTEOME RES . . . . . . 1 . . 1 1 1 . . . . . . . . . . . . . . . . . . . . . . . .  
## YANG Q, 2018, FRONT MICROBIOL . . . . . . . . . . . . 1 1 1 1 1 . . . . . . . . . . . . . . . . . . .  
## CHEN W, 2018, MEDICINE . . . . . . . . . . . . . . . . . 1 1 1 . . . . . . . . . . . . . . . .  
## QIU WR, 2018, GENOMICS . . . . . . . . . . . . . . . . . 1 . . 1 1 1 1 1 . . . . . . . . . . .  
## HUANG H, 2018, PROTEOMICS . . . . . . . . . . . . . . . 1 . 1 . . . . . . . 1 1 1 . . . . . . . .  
## LIU LJ, 2018, J SHANGHAI JIAOTONG UNIV MED SCI . . . . . . . . . . . . . . . . . 1 . . . . . . . . . . 1 1 1 . . . . .  
## BOS J, 2018, J AM CHEM SOC . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## JIANG G, 2018, J CLIN INVEST . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## SUN H, 2017, SCI REP . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## QIU WR, 2017, ARTIF INTELL MED . . . . . . . . . . . . . . . . . . . . . . . . 1 . . . . . . 1 1 1 . .  
## WU Q, 2017, J PROTEOME RES . . . . . . . . . . . 1 . . . . . 1 . . . . . . . . . . . . . . . . 1 1  
## JU Z, 2017, J MOL GRAPH MODEL . . . . . . 1 . . . . . . . . 1 . . . . . . . . . . . . . . . . . 1 . .  
## LIU S, 2017, MOL CELL 1 . . 1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## XU W, 2017, CELL RES . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## WEI W, 2017, CELL RES . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## WEI W, 2017, J PROTEOME RES . . . . . . . . . . . . . . . . . 1 . 1 . . . . . . . . . . . . . . . .  
## LI Y, 2017, TRANSCRIPTION . . . 1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## XIONG X, 2016, NAT CHEM BIOL . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## RUIZ-ANDRES O, 2016, DNM DIS MODELS MECH . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## ANDREWS FH, 2016, NAT CHEM BIOL . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## ZHAO D, 2016, CELL RES . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## LI Y, 2016, MOL CELL . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## SABARI BR, 2015, MOL CELL . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## SABARI BR, 2015, MOL CELL-a . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## BAUMANN K, 2015, NAT REV MOL CELL BIOL . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## MONTELLIER E, 2012, BIOESSAYS . . . . . . . . . . 1 . . . . . . . . . . . . . . . . . . . . . . . . .  
## TAN M, 2011, CELL . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
##   
## CHAN JC, 2019, BIOL PSYCHIATRY . . . . . . . . . . . . . . . . . . . . . .  
## LIU Y, 2019, BIOL PSYCHIATRY . . . . . . . . . . . . . . . . . . . . . .  
## WAN J, 2019, BIOMED PHARMACOTHER . . . . . . . . . . . . . . . . . . . . . .  
## KELLY RDW, 2018, SCI REP . . . . . . . . . . . . . . . . . . . . . .  
## LIU K, 2018, SCI REP . . . . . . . . . . . . . . . . . . . . . .  
## KWON OK, 2018, SCI REP . . . . . . . . . . . . . . . . . . . . . .  
## FELLOWS R, 2018, NAT COMMUN . . . . . . . . . . . . . . . . . . . . . .  
## LIU S, 2018, MOL CELL PROTEOMICS . . . . . . . . . . . . . . . . . . . . . .  
## LU Y, 2018, GENOME BIOL . . . . . . . . . . . . . . . . . . . . . .  
## WU Q, 2018, J PROTEOME RES . . . . . . . . . . . . . . . . . . . . . .  
## YANG Q, 2018, FRONT MICROBIOL . . . . . . . . . . . . . . . . . . . . . .  
## CHEN W, 2018, MEDICINE . . . . . . . . . . . . . . . . . . . . . .  
## QIU WR, 2018, GENOMICS . . . . . . . . . . . . . . . . . . . . . .  
## HUANG H, 2018, PROTEOMICS . . . . . . . . . . . . . . . . . . . . . .  
## LIU LJ, 2018, J SHANGHAI JIAOTONG UNIV MED SCI . . . . . . . . . . . . . . . . . . . . . .  
## BOS J, 2018, J AM CHEM SOC . . . . . . . . . . . . . . . . . . . . . .  
## JIANG G, 2018, J CLIN INVEST . . . . . . . . . . . . . . . . . . . . . .  
## SUN H, 2017, SCI REP . . . . . . . . . . . . . . . . . . . . . .  
## QIU WR, 2017, ARTIF INTELL MED . . . . . . . . . . . . . . . . . . . . . .  
## WU Q, 2017, J PROTEOME RES . . . . . . . . . . . . . . . . . . . . . .  
## JU Z, 2017, J MOL GRAPH MODEL 1 1 . . . . . . . . . . . . . . . . . . . .  
## LIU S, 2017, MOL CELL . . 1 1 1 1 . . . . . . . . . . . . . . . .  
## XU W, 2017, CELL RES . . . . . . . . . . . . . . . . . . . . . .  
## WEI W, 2017, CELL RES . . . . . . . . . . . . . . . . . . . . . .  
## WEI W, 2017, J PROTEOME RES . . . . . . 1 1 1 . . . . . . . . . . . . .  
## LI Y, 2017, TRANSCRIPTION . . . . . . . . . 1 1 1 1 . . . . . . . . .  
## XIONG X, 2016, NAT CHEM BIOL . . . . . . . . . . . . . . . . . . . . . .  
## RUIZ-ANDRES O, 2016, DNM DIS MODELS MECH . . . . . . . . . . . . . 1 1 1 1 1 . . . .  
## ANDREWS FH, 2016, NAT CHEM BIOL . . . . . . . . . . . . . . . . . . . . . .  
## ZHAO D, 2016, CELL RES . . . . . . . . . . . . . . . . . . . . . .  
## LI Y, 2016, MOL CELL . . . . . . . . . . . . . . . . . . . . . .  
## SABARI BR, 2015, MOL CELL . . . . . . . . . . . . . . . . . . . . . .  
## SABARI BR, 2015, MOL CELL-a . . . . . . . . . . . . . . . . . . . . . .  
## BAUMANN K, 2015, NAT REV MOL CELL BIOL . . . . . . . . . . . . . . . . . . . . . .  
## MONTELLIER E, 2012, BIOESSAYS . . . . . . . . . . . . . . . . . . 1 1 1 1  
## TAN M, 2011, CELL . . . . . . . . . . . . . . . . . . . . . .

# Rede de mais Palavras-chave:  
G <- cocMatrix(M, Field = "ID", sep = ";")

# Rede de Países dos Autores:  
N <- metaTagExtraction(M, Field="AU\_CO", sep=";")  
Z <- cocMatrix(M, Field="Au\_CO", sep = ";")

## [1] "Field Au\_CO is not a column name of input data frame"

Obs: Este último não está funcionando.

Entre outros.

#### b. Acoplamento bibliográfico

Para calcular redes de acoplamentos a partir de determinado dado, como Autores, Fontes e Países, utilizamos a função “biblioNetwork”, que usa um argumento de análise e um argumento de rede.

NetMatrix <- biblioNetwork(M, analysis = "coupling", network = "references", sep = ". ")  
NetMatrix

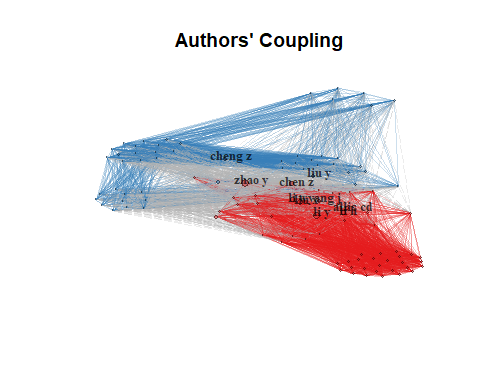
## 36 x 36 sparse Matrix of class "dgCMatrix"

## [[ suppressing 36 column names 'CHAN JC, 2019, BIOL PSYCHIATRY. 2019', 'LIU Y, 2019, BIOL PSYCHIATRY. 2019', 'WAN J, 2019, BIOMED PHARMACOTHER. 2019' ... ]]

##   
## CHAN JC, 2019, BIOL PSYCHIATRY. 2019 1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## LIU Y, 2019, BIOL PSYCHIATRY. 2019 . 1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## WAN J, 2019, BIOMED PHARMACOTHER. 2019 . . 1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## KELLY RDW, 2018, SCI REP. 2018 . . . 1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## LIU K, 2018, SCI REP. 2018 . . . . 1 . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## KWON OK, 2018, SCI REP. 2018 . . . . . 1 . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## FELLOWS R, 2018, NAT COMMUN. 2018 . . . . . . 1 . . . . . . . . . . . . . . . . . . . . . . . . . .  
## LIU S, 2018, MOL CELL PROTEOMICS. 2018 . . . . . . . 1 . . . . . . . . . . . . . . . . . . . . . . . . .  
## LU Y, 2018, GENOME BIOL. 2018 . . . . . . . . 1 . . . . . . . . . . . . . . . . . . . . . . . .  
## WU Q, 2018, J PROTEOME RES. 2018 . . . . . . . . . 1 . . . . . . . . . . . . . . . . . . . . . . .  
## YANG Q, 2018, FRONT MICROBIOL. 2018 . . . . . . . . . . 1 . . . . . . . . . . . . . . . . . . . . . .  
## CHEN W, 2018, MEDICINE. 2018 . . . . . . . . . . . 1 . . . . . . . . 1 . . . . . . . . . . . .  
## QIU WR, 2018, GENOMICS. 2018 . . . . . . . . . . . . 1 . . . . . . . . . . . . . . . . . . . .  
## HUANG H, 2018, PROTEOMICS. 2018 . . . . . . . . . . . . . 1 . . . . . . . . . . . . . . . . . . .  
## LIU LJ, 2018, J SHANGHAI JIAOTONG UNIV MED SCI. 2018 . . . . . . . . . . . . . . 1 . . . . . . . . . . . . . . . . . .  
## BOS J, 2018, J AM CHEM SOC. 2018 . . . . . . . . . . . . . . . 1 . . . . . . . . . . . . . . . . .  
## JIANG G, 2018, J CLIN INVEST. 2018 . . . . . . . . . . . . . . . . 1 . . . . . . . . . . . . . . . .  
## SUN H, 2017, SCI REP. 2017 . . . . . . . . . . . . . . . . . 1 . . . . . . . 1 . . . . . . .  
## QIU WR, 2017, ARTIF INTELL MED. 2017 . . . . . . . . . . . . . . . . . . 1 . . . . . . . . . . . . . .  
## WU Q, 2017, J PROTEOME RES. 2017 . . . . . . . . . . . . . . . . . . . 1 . . . . . . . . . . . . .  
## JU Z, 2017, J MOL GRAPH MODEL. 2017 . . . . . . . . . . . 1 . . . . . . . . 1 . . . . . . . . . . . .  
## LIU S, 2017, MOL CELL. 2017 . . . . . . . . . . . . . . . . . . . . . 1 . . . . . . . . . . .  
## XU W, 2017, CELL RES. 2017 . . . . . . . . . . . . . . . . . . . . . . 1 . . . . . . 1 . . .  
## WEI W, 2017, CELL RES. 2017 . . . . . . . . . . . . . . . . . . . . . . . 1 . . . . . . . . .  
## WEI W, 2017, J PROTEOME RES. 2017 . . . . . . . . . . . . . . . . . . . . . . . . 1 . . . . . . . .  
## LI Y, 2017, TRANSCRIPTION. 2017 . . . . . . . . . . . . . . . . . 1 . . . . . . . 1 . . . . . . .  
## XIONG X, 2016, NAT CHEM BIOL. 2016 . . . . . . . . . . . . . . . . . . . . . . . . . . 1 . . . . . .  
## RUIZ-ANDRES O, 2016, DNM DIS MODELS MECH. 2016 . . . . . . . . . . . . . . . . . . . . . . . . . . . 1 . . . . .  
## ANDREWS FH, 2016, NAT CHEM BIOL. 2016 . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1 . . . .  
## ZHAO D, 2016, CELL RES. 2016 . . . . . . . . . . . . . . . . . . . . . . 1 . . . . . . 1 . . .  
## LI Y, 2016, MOL CELL. 2016 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1 . .  
## SABARI BR, 2015, MOL CELL. 2015 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## SABARI BR, 2015, MOL CELL-a. 2015 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1  
## BAUMANN K, 2015, NAT REV MOL CELL BIOL. 2015 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## MONTELLIER E, 2012, BIOESSAYS. 2012 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## TAN M, 2011, CELL. 2011 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
##   
## CHAN JC, 2019, BIOL PSYCHIATRY. 2019 . . .  
## LIU Y, 2019, BIOL PSYCHIATRY. 2019 . . .  
## WAN J, 2019, BIOMED PHARMACOTHER. 2019 . . .  
## KELLY RDW, 2018, SCI REP. 2018 . . .  
## LIU K, 2018, SCI REP. 2018 . . .  
## KWON OK, 2018, SCI REP. 2018 . . .  
## FELLOWS R, 2018, NAT COMMUN. 2018 . . .  
## LIU S, 2018, MOL CELL PROTEOMICS. 2018 . . .  
## LU Y, 2018, GENOME BIOL. 2018 . . .  
## WU Q, 2018, J PROTEOME RES. 2018 . . .  
## YANG Q, 2018, FRONT MICROBIOL. 2018 . . .  
## CHEN W, 2018, MEDICINE. 2018 . . .  
## QIU WR, 2018, GENOMICS. 2018 . . .  
## HUANG H, 2018, PROTEOMICS. 2018 . . .  
## LIU LJ, 2018, J SHANGHAI JIAOTONG UNIV MED SCI. 2018 . . .  
## BOS J, 2018, J AM CHEM SOC. 2018 . . .  
## JIANG G, 2018, J CLIN INVEST. 2018 . . .  
## SUN H, 2017, SCI REP. 2017 . . .  
## QIU WR, 2017, ARTIF INTELL MED. 2017 . . .  
## WU Q, 2017, J PROTEOME RES. 2017 . . .  
## JU Z, 2017, J MOL GRAPH MODEL. 2017 . . .  
## LIU S, 2017, MOL CELL. 2017 . . .  
## XU W, 2017, CELL RES. 2017 . . .  
## WEI W, 2017, CELL RES. 2017 . . .  
## WEI W, 2017, J PROTEOME RES. 2017 . . .  
## LI Y, 2017, TRANSCRIPTION. 2017 . . .  
## XIONG X, 2016, NAT CHEM BIOL. 2016 . . .  
## RUIZ-ANDRES O, 2016, DNM DIS MODELS MECH. 2016 . . .  
## ANDREWS FH, 2016, NAT CHEM BIOL. 2016 . . .  
## ZHAO D, 2016, CELL RES. 2016 . . .  
## LI Y, 2016, MOL CELL. 2016 . . .  
## SABARI BR, 2015, MOL CELL. 2015 . . .  
## SABARI BR, 2015, MOL CELL-a. 2015 . . .  
## BAUMANN K, 2015, NAT REV MOL CELL BIOL. 2015 1 . .  
## MONTELLIER E, 2012, BIOESSAYS. 2012 . 1 .  
## TAN M, 2011, CELL. 2011 . . 1

Neste exemplo, o acoplamento foi feito de acordo com as referências. Quanto menos referências o artigo tiver, mais fracamente será o acoplamento bibliográfico. Isso sugere que seria mais prático mudar para uma medida relativa. A função “normalizeSimilarity” (correspondente ao argumento “normalize” da função “networkPlot”) calcula a força de associação, inclusão, similaridade de Jaccard ou Salton entre os vértices de uma rede.

NetMatrix1 <- biblioNetwork(M, analysis = "coupling", network = "authors", sep = ";")  
  
net=networkPlot(NetMatrix1, normalize = "salton", weighted=NULL, n = 100, Title = "Authors' Coupling", type = "fruchterman", size=5,size.cex=T,remove.multiple=TRUE,labelsize=0.8,label.n=10,label.cex=F)



#### c. Co-citação bibliográfica

A co-citação ocorre quando dois artigos são citados em um mesmo terceiro artigo. Podemos calculá-la usando a função “biblioNetwork”:

NetMatrix2 <- biblioNetwork(M, analysis = "co-citation", network = "references", sep = ". ")  
NetMatrix2

## 32 x 32 sparse Matrix of class "dgCMatrix"

## [[ suppressing 32 column names 'SUN H. 2013', 'KESSLER R.C. ', 'RUTHENBURG A.J. 2007' ... ]]

##   
## SUN H. 2013 1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## KESSLER R.C. . 1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## RUTHENBURG A.J. 2007 . . 1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## TESSARZ P. 2014 . . . 1 . . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## WALSH C.T. 2005 . . . . 1 . . . . . . . . . . . . . . . . . . . . . . . . . . .  
## MARSHALL R.A. 2016 . . . . . 1 . . . . . . . . . . . . . . . . . . . . . . . . . .  
## KEATING S.T. 2015 . . . . . . 1 . . . . . . . . . . . . . . . . . . . . . . . . .  
## DERIBE Y.L. 2010 . . . . . . . 1 . . . . . . . . . . . . . . . . . . . . . . . .  
## SERVET C. 2010 . . . . . . . . 1 . . . . . . . . . . . . . . . . . . . . . . .  
## CHOUDHARY C. 2014 . . . . . . . . . 1 . . . . . . . . . . . . . . . . . . . . . .  
## ARAFAT W. 1985 . . . . . . . . . . 1 . . . . . . . . . . . . . . . . . . . . .  
## KOUZARIDES T. 2007 . . . . . . . . . . . 2 . . . . . . . . . . . . . . . . . . . .  
## BERDASCO M. 2010 . . . . . . . . . . . . 1 . . . . . . . . . . . . . . . . . . .  
## ECKBURG P.B. 2005 . . . . . . . . . . . . . 1 . . . . . . . . . . . . . . . . . .  
## SIEGEL R.L. 2017 . . . . . . . . . . . . . . 1 . . . . . . . . . . . . . . . . .  
## CHEN T. 2014 . . . . . . . . . . . . . . . 1 . . . . . . . . . . . . . . . .  
## BRUNER K.M. 2015 . . . . . . . . . . . . . . . . 1 . . . . . . . . . . . . . . .  
## ALLFREY V.G. 1964 . . . . . . . . . . . . . . . . . 2 . . . . . . . . . . . . . .  
## TAN M. 2011-1 . . . . . . . . . . . . . . . . . . 1 . . . . . . . . . . . . .  
## TAN M. 2011-2 . . . . . . . . . . . . . . . . . . . 1 . . . . . . . . . . . .  
## AGNIHOTRI G. . . . . . . . . . . . . . . . . . . . . 1 . . . . . . . . . . .  
## TAN M. 2011-3 . . . . . . . . . . . . . . . . . . . . . 2 . . . . . . . . . .  
## ROTH S.Y. 2001 . . . . . . . . . . . . . . . . . . . . . . 1 . . . . . . . . .  
## ESKELAND R. 2010 . . . . . . . . . . . . . . . . . . . . . . . 1 . . . . . . . .  
## STRAHL B.D. 2000 . . . . . . . . . . . . . . . . . . . . . . . . 1 . . . . . . .  
## BAEK J.-H. 2015 . . . . . . . . . . . . . . . . . . . . . . . . . 1 . . . . . .  
## TAN M. 2011-4 . . . . . . . . . . . . . . . . . . . . . . . . . . 1 . . . . .  
## BAO X. 2014 . . . . . . . . . . . . . . . . . . . . . . . . . . . 1 . . . .  
## AN W. 2002 . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1 . . .  
## SABARI B.R. 2015 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1 . .  
## BOUSSOUAR F. 2008 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1 .  
## ALLIS C.D. 1985 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1

#### d. Colaboração bibliográfica

A rede de colaboração científica é uma rede na qual os pontos são os autores e as linhas que os ligam são coautorias. Podemos calculá-la usando a função “biblioNetwork”:

NetMatrix3 <- biblioNetwork(M, analysis = "collaboration", network = "authors", sep = ";")

Para uma rede de colaboração do país seria:

# NetMatrix4 <- biblioNetwork(M, analysis="collaboration", network="countries", sep=";")

Obs: Não está funcionando

### 10. Análise descritiva das características gráficas de rede

A função “networkStat” calcula diferentes estatísticas partindo de uma matriz bibliográfica ou objeto igraph, computando o resumo estatístico da rede e os principais índices de centralidade e prestígio das vértices. Por exemplo:

NetMatrix5 <- biblioNetwork(M, analysis = "co-occurrences", network = "keywords", sep = ";")  
netstat <- networkStat(NetMatrix)  
netstat

## $graph  
## IGRAPH e1dd3f0 UN-- 36 3 --   
## + attr: name (v/c), id (v/c)  
## + edges from e1dd3f0 (vertex names):  
## [1] CHEN W, 2018, MEDICINE. 2018--JU Z, 2017, J MOL GRAPH MODEL. 2017  
## [2] SUN H, 2017, SCI REP. 2017 --LI Y, 2017, TRANSCRIPTION. 2017   
## [3] XU W, 2017, CELL RES. 2017 --ZHAO D, 2016, CELL RES. 2016   
##   
## $network  
## $network$networkSize  
## [1] 36  
##   
## $network$networkDensity  
## [1] 0.004761905  
##   
## $network$networkTransitivity  
## [1] NaN  
##   
## $network$networkDiameter  
## [1] 1  
##   
## $network$networkDegreeDist  
## [1] 1.0000000 0.1666667  
##   
## $network$networkCentrDegree  
## [1] 0.02380952  
##   
## $network$networkCentrCloseness  
## [1] NA  
##   
## $network$networkCentrEigen  
## [1] NA  
##   
## $network$networkCentrbetweenness  
## [1] NA  
##   
## $network$NetworkAverPathLeng  
## [1] 1  
##   
##   
## $vertex  
## [1] NA  
##   
## $stat  
## [1] "network"  
##   
## $type  
## [1] "degree"  
##   
## attr(,"class")  
## [1] "bibliometrix\_netstat"

O resumo estatístico da rede fornecido possui:

names(netstat$network)

## [1] "networkSize" "networkDensity" "networkTransitivity" "networkDiameter"   
## [5] "networkDegreeDist" "networkCentrDegree" "networkCentrCloseness" "networkCentrEigen"   
## [9] "networkCentrbetweenness" "NetworkAverPathLeng"

Porém não é possível listar os principais índices de centralidade e prestígio das vértices:

names(netstat$vertex)

## NULL

Para resumir os principais resultados da função “networkStat”, usamos “summary”:

summary(netstat, k=10)

##   
##   
## Main statistics about the network  
##   
## Size 36   
## Density 0.005   
## Transitivity NaN   
## Diameter 1   
## Degree Centralization 0.024   
## Average path length 1   
##

Ele exibe as principais informações sobre a rede e a descrição do vértice através de várias tabelas.

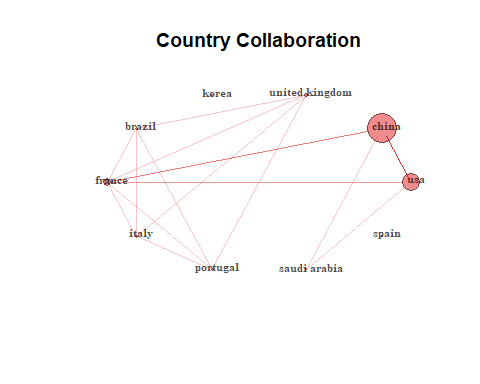
### 11. Visualizando redes bibliográficas

Podemos visualizar ou modeladar graficamente todas as redes bibliográficas utilizando a função “networkPlot”.Seu argumento principal é type, que indica o layout do mapa de rede: circle, kamada-kawai, mds, entre outros.

#### a. Países com Colaboração Científica

Para criar uma rede de países colaboradores e plotar um gráfico, utilizamos os seguintes códigos:

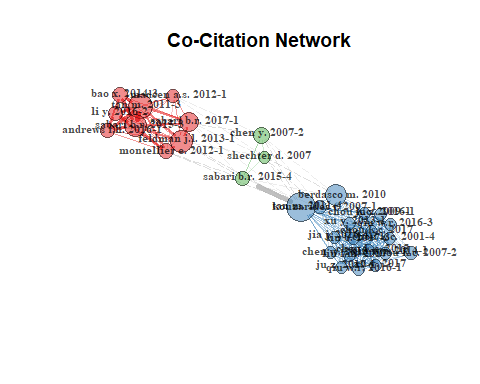
S <- metaTagExtraction(M, Field = "AU\_CO", sep = ";")  
NetMatrix6 <- biblioNetwork(S, analysis = "collaboration", network = "countries", sep = ";")  
  
net=networkPlot(NetMatrix6, n = dim(NetMatrix6)[1], Title = "Country Collaboration", type = "circle", size=TRUE, remove.multiple=FALSE,labelsize=0.7,cluster="none")



#### b. Rede de co-citação

Para criar uma rede de co-citação e plotar um gráfico, usamos:

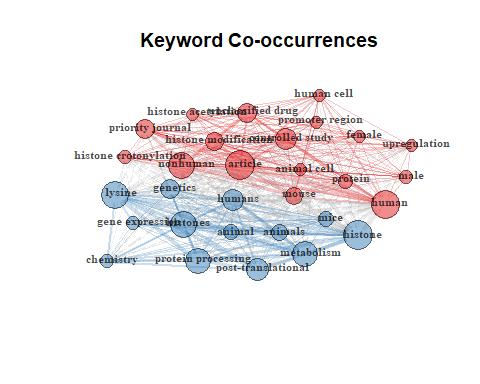
NetMatrix7 <- biblioNetwork(M, analysis = "co-citation", network = "references", sep = ";")  
  
net=networkPlot(NetMatrix7, n = 30, Title = "Co-Citation Network", type = "fruchterman", size=T, remove.multiple=FALSE, labelsize=0.7,edgesize = 5)



#### c. Co-ocorrências de palavras-chave

Para criar uma rede de co-ocorrências de palavras-chave e plotar um gráfico seria por:

NetMatrix8 <- biblioNetwork(M, analysis = "co-occurrences", network = "keywords", sep = ";")  
  
net=networkPlot(NetMatrix8, normalize="association", weighted=T, n = 30, Title = "Keyword Co-occurrences", type = "fruchterman", size=T,edgesize = 5,labelsize=0.7)

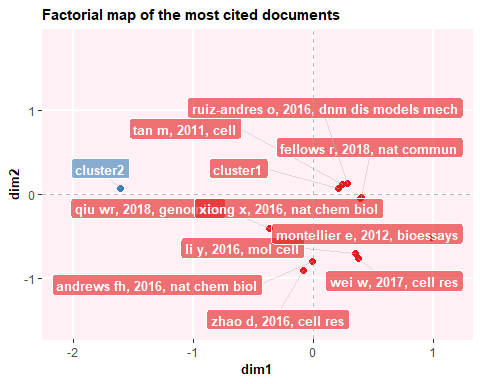
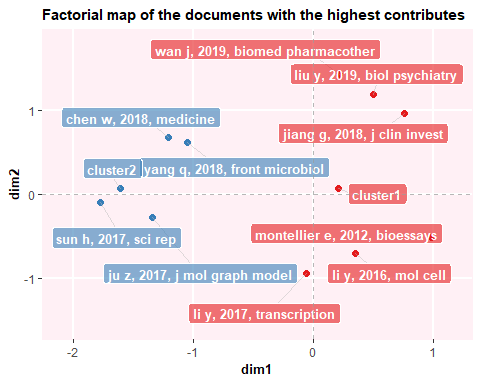
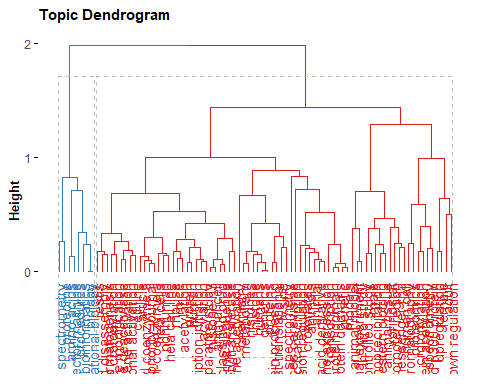
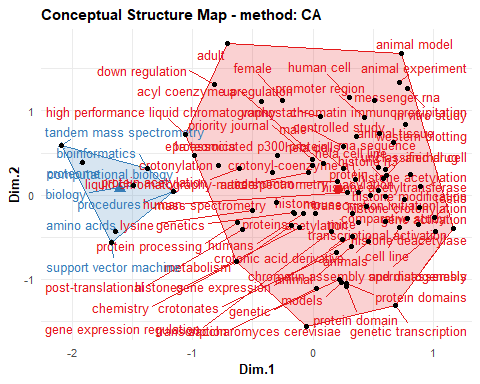


### 12. Análise Conjunta: A estrutura conceitual de um campo

O objetivo dessa análise é mapear a estrutura conceitual de um framework usando co-ocorrências em uma coleção bibliográfica. Pode ser feita por Escala Multidimensional (MDS), Análise de Correspondência (AC) ou Análise de Correspondência Múltipla (ACM).

A função “conceptualStructure” extrai termos de títulos e resumos, além de reduzir as palavras flexionadas ou derivadas para as palavras originais. Um exemplo utilizando palavras-chave:

CSK <- conceptualStructure(M,field="ID", method="CA", minDegree=4, k.max=8, stemming=FALSE, labelsize=10, documents=10)



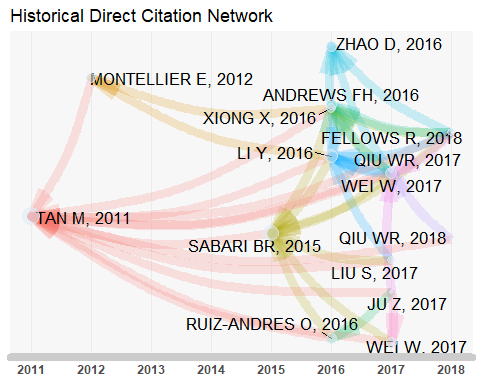
### 13. Rede Histórica de Citação Direta

O mapa historiográfico representa um mapa das citações diretas mais relevantes resultantes de uma coleção bibliográfica em rede cronológica, podendo ser plotada usando “histPlot”:

options(width=130)  
histResults <- histNetwork(M, min.citations = 10, sep = ";")

## Articles analysed 15

net <- histPlot(histResults, n=15, size = 20, labelsize=10, size.cex=TRUE, arrowsize = 0.5, color = TRUE)



##   
## Legend  
##   
## Paper DOI Year LCS GCS  
## 2011 - 1 TAN M, 2011, CELL 10.1016/J.CELL.2011.08.008 2011 28 729  
## 2012 - 2 MONTELLIER E, 2012, BIOESSAYS 10.1002/BIES.201100141 2012 5 54  
## 2015 - 3 SABARI BR, 2015, MOL CELL-a 10.1016/J.MOLCEL.2015.02.029 2015 21 121  
## 2016 - 4 XIONG X, 2016, NAT CHEM BIOL 10.1038/NCHEMBIO.2218 2016 5 32  
## 2016 - 5 RUIZ-ANDRES O, 2016, DNM DIS MODELS MECH 10.1242/DMM.024455 2016 7 20  
## 2016 - 6 ANDREWS FH, 2016, NAT CHEM BIOL 10.1038/NCHEMBIO.2065 2016 7 49  
## 2016 - 7 ZHAO D, 2016, CELL RES 10.1038/CR.2016.49 2016 8 33  
## 2016 - 8 LI Y, 2016, MOL CELL 10.1016/J.MOLCEL.2016.03.028 2016 15 71  
## 2017 - 9 QIU WR, 2017, ARTIF INTELL MED 10.1016/J.ARTMED.2017.02.007 2017 1 10  
## 2017 - 10 JU Z, 2017, J MOL GRAPH MODEL 10.1016/J.JMGM.2017.08.020 2017 0 13  
## 2017 - 11 LIU S, 2017, MOL CELL 10.1016/J.MOLCEL.2017.07.011 2017 3 13  
## 2017 - 12 WEI W, 2017, CELL RES 10.1038/CR.2017.68 2017 10 21  
## 2017 - 13 WEI W, 2017, J PROTEOME RES 10.1021/ACS.JPROTEOME.7B00012 2017 8 16  
## 2018 - 14 FELLOWS R, 2018, NAT COMMUN 10.1038/S41467-017-02651-5 2018 3 26  
## 2018 - 15 QIU WR, 2018, GENOMICS 10.1016/J.YGENO.2017.10.008 2018 0 42

## Respostas encontradas

Como mostrado em cada tópico, o pacote “bibliometrix” permite a análise bibliométrica com criação de gráficos, redes, tabelas e rankings de acordo com parâmetros importantes no meio científico, como quantidade de citações, produtividade e colaborações. Alguns dos resultados obtidos na análise de artigos que possuem crotonilação (*crotonylation* em inglês) em seu título mostram que foram produzidos entre 2011 e 2019, principalmente na China, sendo o autor mais produtivo Lu L. e o mais citado Zhao Y. Essas informações são interessantes quando pensamos em fazer um novo projeto, revisão bibliográfica ou até mesmo tentar estabelecer contato com algum pesquisador da área, focando naqueles que possuem melhores avaliações.

## Dificuldades encontradas

As principais dificuldades que tive estão relacionadas ao uso do Rmarkdown e GitHub. No Rmarkdown, por exemplo, não consegui diminuir o tamanho da fonte dos gráficos, o que acabou dificultando a visualização de alguns dados pela sobreposição de escritos. Além disso, os códigos referentes a Rede de Países dos Autores (N e Z) no tópico 9.a. e a Rede de Colaboração do País (NetMatrix4) no tópico 9.d. não funcionaram e não soube como consertá-los, mesmo pedindo auxílio a colegas de classe. Quanto ao GitHub, todo o processo para conseguir enviar meus arquivos a um diretório foi complicada. Compreendi melhor buscando tutoriais, que se encontram na bibliografia.

## Bibliografia

#### Rmarkdown:

* Writing documents with R Markdown (<https://monashbioinformaticsplatform.github.io/2017-11-16-open-science-training/topics/rmarkdown.html>)
* Manual do pacote ‘rmarkdown’ (<https://cran.r-project.org/web/packages/rmarkdown/rmarkdown.pdf>)

#### Git:

* RStudio e GitHub no dia a dia (<https://www.curso-r.com/blog/2017-07-17-rstudio-e-github/>)
* Git and GitHub (<http://r-pkgs.had.co.nz/git.html>)

#### Bibliometrix:

* A brief introduction to bibliometrix (<https://cran.r-project.org/web/packages/bibliometrix/vignettes/bibliometrix-vignette.html>)
* Manual do pacote ‘bibliometrix’ (<https://cran.r-project.org/web/packages/bibliometrix/bibliometrix.pdf>)

#### Documento Scopus.bib:

* SCOPUS (<http://www.scopus.com>)