PAGnet: *Pseudomonas aeruginosa* genomic integrated regulatory network.

***Hao Huang*[1], *Xiaolong Shao*[2], *Yingpeng Xie*[1], *Tingting Wang*[1], *Yingchao Zhang*[2], *Xin Wang*[1], *and Xin Deng*[1]**

[1]Department of Biomedical Sciences, City University of Hong Kong, Hong Kong
[2]Key Laboratory of Molecular Microbiology and Technology, Ministry of Education, TEDA Institute of Biological Sciences and Biotechnology, Nankai University, 23 Hongda Street, Tianjin 300457

2019-03-08

Abstract

R package PAGnet is an R package to provide the user with *Pseudomonas aeruginosa* genomic integrated regulatory network visualization, subnetwork filtering and master regulator analysis in his/her comoputer, consistent with the functionality which deployed on the online website (http://pagnetwork.org). The package could facilitate visualization and exploration of regulatory network, as well as master regulator analysis for identification of key transcription factors mediating a biological process or pathway.

Package

PAGnet 0.1.0

Contents

1 Overview

Regulatory networks including virulence-related transcriptional factors (TFs) determine bacterial pathogenicity in response to different environmental cues. *Pseudomonas aeruginosa*, a Gram-negative opportunistic pathogen of humans, recruits numerous TFs in quorum sensing (QS) system, type III secretion system (T3SS) and Type VI secretion system (T6SS) to mediate the pathogenicity. Although many virulence-related TFs have been illustrated individually, very little is known about their crosstalks and regulatory network. Here, based on chromatin immunoprecipitation coupled with high-throughput sequencing (ChIP-seq) and transcriptome profiling (RNA-seq), we primarily focused on understanding the crosstalks of 20 virulence-related TFs, which led to construction of a virulence regulatory network named PAGnet (*Pseudomonas aeruginosa* genomic integrated regulatory network) including 82 crosstalk targets.

The PAGnet uncovered the intricate mechanism of virulence regulation and revealed master regulators in QS, T3SS and T6SS pathways. The package **PAGnet** is designed for network visualization, subnetwork filtering and master regulator analysis locally by running a local shniy GUI within R. In addition, the master regulator analysis can be performed in the R console without runnning the shiny GUI to provide more flexibility.

Before starting to use this package, you need to install the following packages:

```
library(shiny)
library(shinythemes)
library(scales)
```

Then load the package:

```
  ## loading package

  library(PAGnet)
```

2 Quick Start

2.1 Master Regulator Analysis

The user can choose to use the default PAGnet or to upload their own regulatory network in a predefined format (Locus Tag).

A dataframe format of default **PAGnet** with two colmn transcription factors and targets is provided for MRA, user could also their own regulatory network in same format as input.

```
  data(PAGnet)

  ##Use defaultPAGnet as regulatory network

  head(pagnet)
##   TranscriptionFactor Target
```

```
## 1              PA5261 PA0510
## 2              PA5261 PA0511
## 3              PA5261 PA0515
## 4              PA5261 PA0527
## 5              PA5261 PA1048
## 6              PA5261 PA1590
```

The users also need to provide a character vector of TFs (or only interested TFs).

```
  ## define transcription factors in regulatory network

  head(tf)
##     AlgR     AmrZ     BfmR     CdpR     ExsA     FleQ
## "PA5261" "PA3385" "PA4101" "PA2588" "PA1713" "PA1097"
```

A character vector of locus tag should be provide as signature genes.

```
  ##Use QS related genes as signature genes

  head(qs)
##     rhlC     rhlG     rhlI     rhlR     rhlA     rhlB
## "PA1130" "PA3387" "PA3476" "PA3477" "PA3479" "PA3478"
```

The function `pagnet.mra` is used to perform MRA over provided regulatory network. The MRA computes the overlap between the transcriptional regulatory unities (regulons) and the input signature genes using the hypergeometric distribution (with multiple hypothesis testing corrections). Having completed master regulator analysis, a table will be returned.

```
  ## Perform MRA and return results

  mra_results <- pagnet.mra(rnet=pagnet,tflist=tf,signature = qs,
                            pValueCutoff = 0.05)

  mra_results
##        TF network.size regulon.size signature.size observed.signature.size
## 14 PA3477          533            5             13                       3
## 9  PA1430          533           49             13                       6
## 15 PA1431          533           12             13                       3
## 11 PA1003          533           19             13                       3
##    Pvalue
## 14  1e-04
## 9   5e-04
## 15 0.0022
## 11 0.0088
```

2.2 Local shiny interface

The function `pagnet.mra.interface` is used to call **local shiny GUI**.

```
## call local shony interfact
pagnet.mra.interface()
```

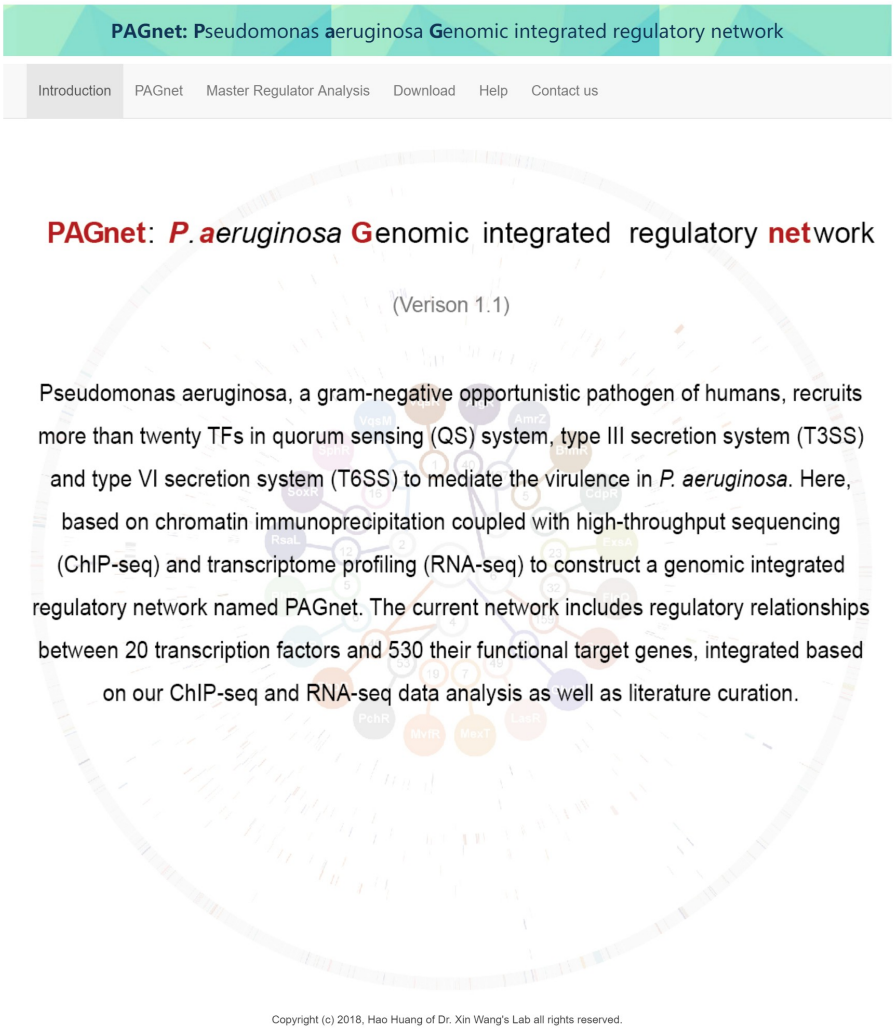

Figure 1: **A screenshot of local shiny interface of PAGnet**

The shiny GUI provides visualization and exploration of PAGnet. It allows the user to filter the full network by selecting one or multiple transcription factor(s) to obtain a subnetwork of interest. A brief summary of the subnetwork is also provided with information about the transcription factors and their target genes.

The local shiny GUI also provides master regulator analysis for identification of key transcription factors mediating a biological process or pathway. First, the user can choose to use the default PAGnet or to upload their own regulatory network in a predefined format. Second, the user needs to specify a gene signature associated with a biological function or pathway of interest, either by selecting a gene set from public databases or uploading a user-customized gene list. In the current version, the platform provides gene sets in Gene Ontology (GO) and KEGG databases obtained from Pseudomonas Genome DB (Winsor et al. (2015)). Having completed master regulator analysis, a table will be returned with information about each transcription

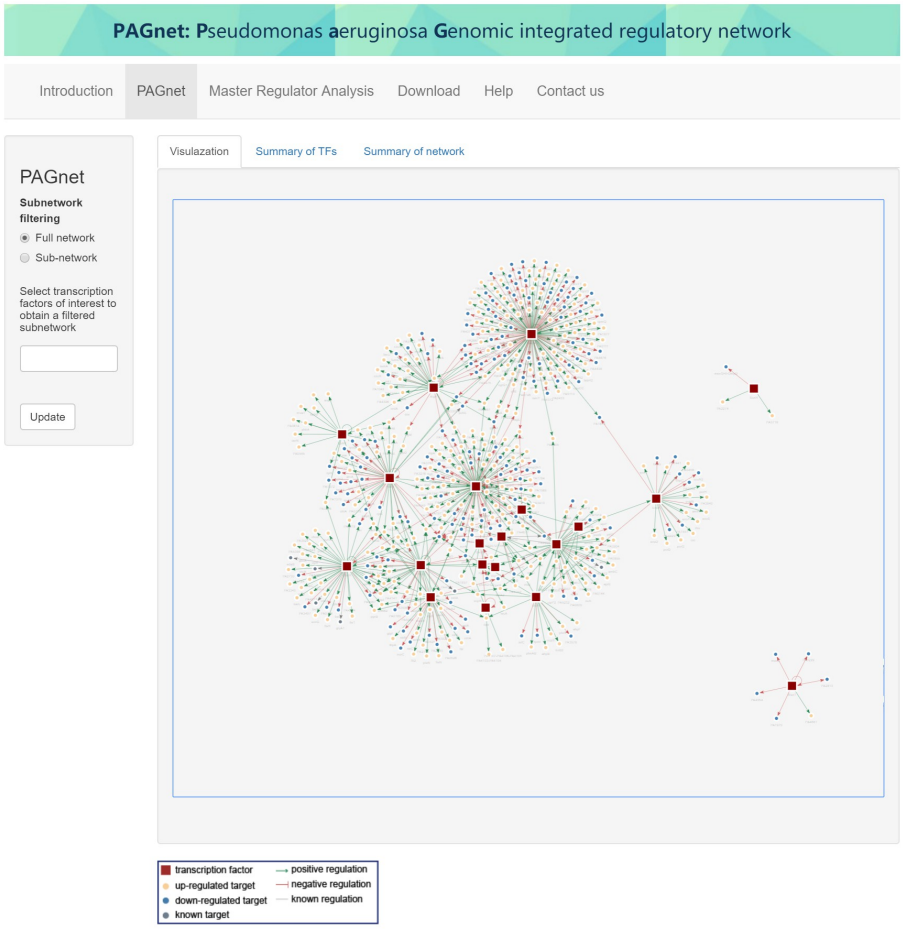

Figure 2: **Visualization of PAGnet in local shiny GUI**

factor's corresponding gene ID, gene name, number of target genes, total number of hits (all signature genes in the network), observed hits (signature genes in the TF's regulon), and a p-value calculated based on a hypergeometric test. The table is sorted according to the statistical significance indicated by the p-values, and the top significant TFs can be prioritized as master regulators.

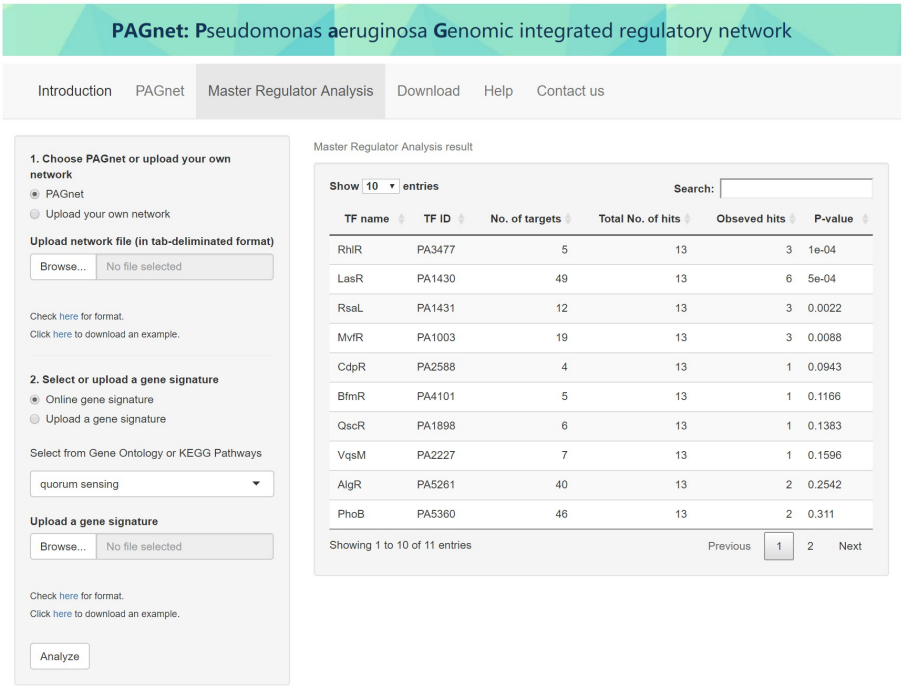

Figure 3: Master Regulator Analysis

3 Need help?

If you have any question/issue, please feel free to contact us.

```
Dr. Xin Deng (xindeng@cityu.edu.hk)
```

or

```
Dr. Xin Wang (xin.wang@cityu.edu.hk)
```

4 Session Information

```
## R version 3.5.2 (2018-12-20)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Ubuntu 18.04.1 LTS
##
## Matrix products: default
## BLAS: /usr/lib/x86_64-linux-gnu/openblas/libblas.so.3
## LAPACK: /usr/lib/x86_64-linux-gnu/libopenblasp-r0.2.20.so
##
## locale:
##  [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
##  [3] LC_TIME=en_US.UTF-8        LC_COLLATE=en_US.UTF-8
```

```
##  [5] LC_MONETARY=en_US.UTF-8    LC_MESSAGES=en_US.UTF-8
##  [7] LC_PAPER=en_US.UTF-8       LC_NAME=C
##  [9] LC_ADDRESS=C               LC_TELEPHONE=C
## [11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] PAGnet_0.1.0      scales_1.0.0      shinythemes_1.1.2 shiny_1.2.0
## [5] BiocStyle_2.10.0
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_0.12.18       knitr_1.21         magrittr_1.5
##  [4] munsell_0.5.0      colorspace_1.3-2   xtable_1.8-3
##  [7] R6_2.3.0           stringr_1.3.1      tools_3.5.2
## [10] xfun_0.4           htmltools_0.3.6    yaml_2.2.0
## [13] digest_0.6.15      bookdown_0.8       BiocManager_1.30.4
## [16] later_0.7.3        promises_1.0.1     evaluate_0.12
## [19] mime_0.6           rmarkdown_1.11     stringi_1.2.3
## [22] compiler_3.5.2     httpuv_1.4.4.1
```

References

Winsor, Geoffrey L, Emma J Griffiths, Raymond Lo, Bhavjinder K Dhillon, Julie A Shay, and Fiona SL Brinkman. 2015. "Enhanced Annotations and Features for Comparing Thousands of Pseudomonas Genomes in the Pseudomonas Genome Database." *Nucleic Acids Research* 44 (D1). Oxford University Press: D646–D653.