TECHNICAL NOTE

GUIdock-VNC: using a graphical desktop sharing system to provide a browser-based interface for containerized software

Varun Mittal, Ling-Hong Hung, Jayant Keswani, Daniel Kristiyanto, Sung Bong Lee and Ka Yee Yeung*

Institute of Technology, University of Washington, Tacoma, 1900 Commerce St, 98402, Tacoma, WA, USA

*Correspondence: kayee@uw.edu

Abstract

Background: Software container technology such as Docker can be used to package and distribute bioinformatics workflows consisting of multiple software implementations and dependencies. However, Docker is a command line–based tool, and many bioinformatics pipelines consist of components that require a graphical user interface. **Results:** We present a container tool called GUIdock-VNC that uses a graphical desktop sharing system to provide a browser-based interface for containerized software. GUIdock-VNC uses the Virtual Network Computing protocol to render the graphics within most commonly used browsers. We also present a minimal image builder that can add our proposed graphical desktop sharing system to any Docker packages, with the end result that any Docker packages can be run using a graphical desktop within a browser. In addition, GUIdock-VNC uses the Oauth2 authentication protocols when deployed on the cloud. **Conclusions:** As a proof-of-concept, we demonstrated the utility of GUIdock-noVNC in gene network inference. We benchmarked our container implementation on various operating systems and showed that our solution creates minimal overhead.

Keywords: research reproducibility; Docker; containers; software; graphical user interface; bioinformatics

Background

Modern workflows in computational fields such as bioinformatics consist of multiple software implementations, each with their own set of dependencies. Software container technology such as Docker (http://www.docker.com) package the dependencies with the software and provide a method to reproduce these complex pipelines on multiple hardware and cloud platforms. For example, BioShadock [1], BioDocker and Bioboxes [2] are two frameworks aimed at reproducibly deploying bioinformatics workflows using Docker containers.

Many bioinformatics pipelines have a component that requires a graphical user interface (GUI) that can potentially limit the portability of the Dockerized workflows as different platforms use different methodologies to render the GUI. We have previously described and implemented GUIdock-X11 [3], an X11-based methodology for portably supporting GUI applications in containers on different platforms. While the X11-based display method can be conveniently deployed in the local environment by exposing a file socket from a container, deploying the image on a cloud and accessing it remotely is non-trivial. In addition, on systems such as Windows, where there is no native X11

support, additional client software must be installed by the user to render the X11 graphics locally. Here we describe GUIdock-VNC, which implements an improved browser-based solution that does not require the user to map ports, configure firewalls, or install any additional specialized software.

GUIdock-VNC uses the Virtual Network Computing (VNC) protocol [4] to render the graphics. Instead of transferring commands and allowing a local client to render the graphics, VNC transfers a pre-rendered screen. Bandwidth requirements are minimized by only transferring the differences between the current screen and the last screen. This can actually be less chatty than the X11 methodology, which is constantly sending display commands. noVNC is a browser-based VNC client implemented using HTML5 Canvas and WebSockets [5]. Modern browsers can use the HTML5-based noVNC client to display the screen locally. The browser transparently downloads the noVNC client from the container and becomes the terminal, thus eliminating the need for the user to configure and install separate software. This is a major advantage as the users of bioinformatics workflows are not necessarily technically trained in configuring computer systems.

Most importantly, GUIdock-VNC also facilitates the deployment of Docker applications on the cloud. With a browser-based solution, we also have access to web-based authentication protocols such as Oauth2 [6], which allows for authentication using an email account. The host service is accessed and authenticated through the HTTP/HTTPS port, greatly simplifying the configuration necessary to support cloud-based platforms.

Our contributions

We implemented GUIdock-VNC, which adds and configures a software layer inside a Docker container to allow applications to export a GUI using the VNC protocol. When deployed on the cloud, authentication is provided using Oauth2. In addition, we provide a set of minimal base images to allow the users to add the host graphical desktop interface to any existing Dockerfiles. No client software installation is necessary for the users as GUIdock-VNC uses the HTML5 noVNC browser-based client to display the GUI. All our tools are publicly available on GitHub.

We benchmarked the implementation on a real-world bioinformatics pipeline. Our results showed that noVNC creates minimal overhead and GUIdock-VNC is superior to our previous work, GUIdock-X11 [3], and other virtual machine–based deployment solutions.

Related work

Software containers and Docker
A software container packages an application with everything it needs to run, including supporting libraries and system resources. Containers differ from traditional virtual machines

(VMs) in that the resources of the operating system (OS), and not the hardware, are virtualized. In addition, multiple containers share a single OS kernel, thus saving considerable resources over multiple VMs.

Linux has supported OS-level virtualization for several years. Docker (http://www.docker.com/) is an open source project that provides tools to setup and deploy Linux software containers. While Docker can run natively on Linux hosts, a small Linux VM is necessary to provide the virtualization services on Mac OS and Windows systems. On non-Linux systems, a single Docker container consists of a mini-VM, the Docker software layer, and the software container. However, multiple Docker containers can share the same mini-VM, saving considerable resources over using multiple individual VMs. Recently, support for OS-level virtualization has been added to Windows and the Macintosh operating system (Mac OS). Beta versions of Docker for both Windows and Mac OS are now available that allow Docker to run natively. Subsequently, these beta versions allow native Windows and Mac OS software to be containerized and deployed in a similar manner [7]. Docker containers therefore provide a convenient and light method for deploying open source workflows on multiple platforms.

GUIdock-X11
Although Docker provides a container with the original software environment, the host system, where the container software is executed, is responsible for rendering graphics. Our previous work, GUIdock-X11 [3], is one of the solutions in bridging the graphical information from user and Docker containers by using the X11 common graphic interface. GUIdock-X11 passes the container X11 commands to a host X11 client, which renders the GUI. Security is handled by encrypting the commands through secure shell (ssh) tunneling. We demonstrated the use of GUIdock-X11 [3] for systems biology applications, including Bioconductor packages written in R, C++, and Fortran, as well as Cytoscape, a standalone Java-based application with a graphical user interface. Neither Windows nor Mac OS uses X11 natively to render their graphics. Additional software such as MobaXterm [8] or socat [9] is needed to emulate X11 and locally render the graphics commands exported by the Docker container. However, a major advantage of the X11 method is that the commands to render the graphics and not the graphics themselves are transmitted, potentially reducing the total bandwidth required.

Table 1 summarizes the differences between GUIdock-VNC and our previous work, GUIdock-X11.

Case study: inference of gene networks
The inference of gene networks is a fundamental challenge in systems biology. We use gene network inference as a case study to demonstrate that GUIdock-X11 and GUIdock-VNC can be used to yield reproducible results from bioinformatics workflows. We have previously developed inference methods

Table 1: Comparison between GUIdock-X11 and GUIdock-VNC

Feature	GUIdock-X11	GUIdock-VNC
Can be deployed on phones/tablets?	No	Yes
Security	ssh-tunnel	OAuth2
Bandwidth	Low	Low to medium
Cloud integration difficulty	Medium	Simple
Dockerfile setup	Manual editing	Automatic conversion of base Docker images
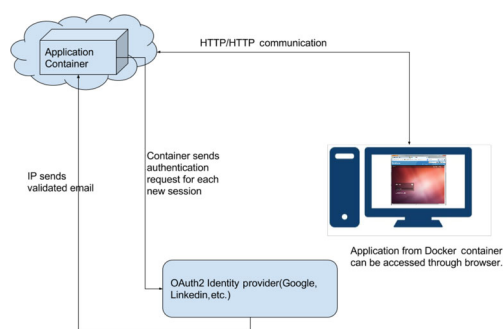
Figure 1: Architectural overview of GUIdock-VNC. In the proposed architecture, each container is a self-contained web server that can be accessed using a single port. When deployed on the cloud, the services can be accessed using the cloud provider's network address translation (NAT) mechanism. Each container is also capable of OAuth2 authentication that can be enabled while deploying the application. Once enabled, the user will be required to sign in through an identity provider (such as Google in the current prototype). After the authentication, the user browser will be automatically redirected to the application.

using a regression-based framework, in which we searched for candidate regulators (i.e., parent nodes) for each target gene [10–12]. Our methods are implemented in R, C++, and Fortran, and the implementation is available as a Bioconductor package called networkBMA (http://bioconductor.org/packages/release/bioc/html/networkBMA.html) [13]. In order to visualize the resulting gene networks, we previously developed a Cytoscape app called CyNetworkBMA (http://apps.cytoscape.org/apps/cynetworkbma) [14]. Cytoscape is a Java-based stand-alone application with a GUI to analyze and visualize graphs and networks [15–17]. Our app, CyNetworkBMA [14], integrates our networkBMA Bioconductor package into Cytoscape, allowing the user to directly visualize the resulting gene networks inferred from networkBMA using the Cytoscape utilities. The integration of multiple pieces of software, each with its own software dependencies, makes CyNetworkBMA an ideal proof-of-concept application for the illustration of the utility of GUIdock-VNC.

Implementation of GUIdock-VNC

Fig. 1 shows an overview of GUIdock-VNC.

Virtual Network Computing

VNC is a framebuffer-based protocol that was written to view and control a remote desktop over the internet [4]. VNC is essentially a server program that attaches to a display server like X11 and creates a proxy between the client and the display server. The proxy server takes in input from the client and relays it to the display server while at the same time the display server sends pre-rendered display images to the client. VNC is thus a network intensive protocol, although the amount of data transferred back and forth can be reduced by using various compression technologies on the transfer layer. In our implementation, we use Xvnc/Xvfb (X virtual frame buffer) [18] to provide a lightweight VNC/X11 display server.

noVNC

noVNC is a browser-based VNC client [5]. The name 'noVNC' means that the traditional VNC client is not needed and that a modern browser with HTML5 and WebSocket support can be used to access and control a remote VNC server. This noVNC technology is particularly interesting as almost all browsers, both desktop and mobile versions, have HTML5 extensions

Docker application image
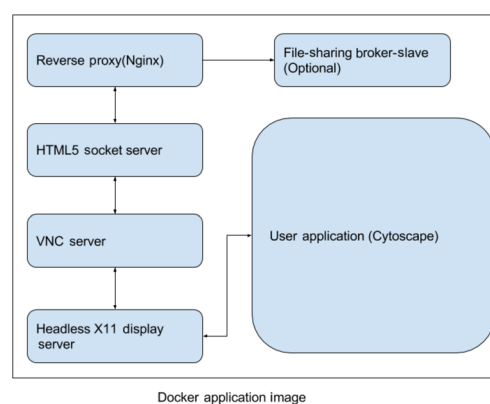
Figure 2: Services running inside container. Apart from user applications, there are two web services running inside the container and a reverse proxy (nginx) to act as an interface for the container. The first web service is the noVNC interface connected to the VNC server. The noVNC server is a Python plus JavaScript application framework used for establishing WebSocket for VNC packet interchange. The second web service is an optional broker service that helps in exchanging data through a datastore interface (Mongodb) and a message-passing queue (RabbitMQ).

built in. An additional layer on the host is required for the VNC server to communicate through WebSocket. We use nginx (https://www.nginx.com/), a fast and light reverse-proxy web server, for this purpose.

Authentication methods: Oauth2

For containers that are not deployed locally, i.e., on a network or cloud, security is a concern as the traffic between the viewer and server can be seen by anyone with access to the network. Oauth2 [6] is an authentication method that is commonly used by major corporations to validate third-party applications. Specifically, Oauth allows users to log onto third-party websites using their existing Google, Twitter, or Facebook accounts, thus avoiding the creation of additional accounts for the third-party websites. In the present era, where it is extremely difficult to host identity services and secure communication, public authentication services like Oauth2 plays an important role. Providers like Google, Facebook, and LinkedIn can be used to validate any user registered with an email at one of these websites. In our prototype, we have created a prototype for the Google identity server. Each container can be forced to login through one of these public providers.

Automatic conversion of base Docker images

To add the noVNC graphical desktop to a Docker image, the converted image requires a web server, a headless display server running inside the container, and a VNC server (see Fig. 2). The web server is required to serve the JavaScript-based noVNC client, and the headless display server routes all drawing instructions to the VNC server; they are then sent to the client running the browser-based noVNC JavaScript client. Due to the three active components running inside the container, we generate a bash script to work as the entry point for the container. Therefore, in order to assist users and to let them start application-dependent services, we have created a tool to bootstrap standard images with noVNC capability. The tool accepts a JSON script as input with defined parameters for files to be copied, additional software to be installed (using apt-get), and the location of the startup script, which is then tied into the entry point.
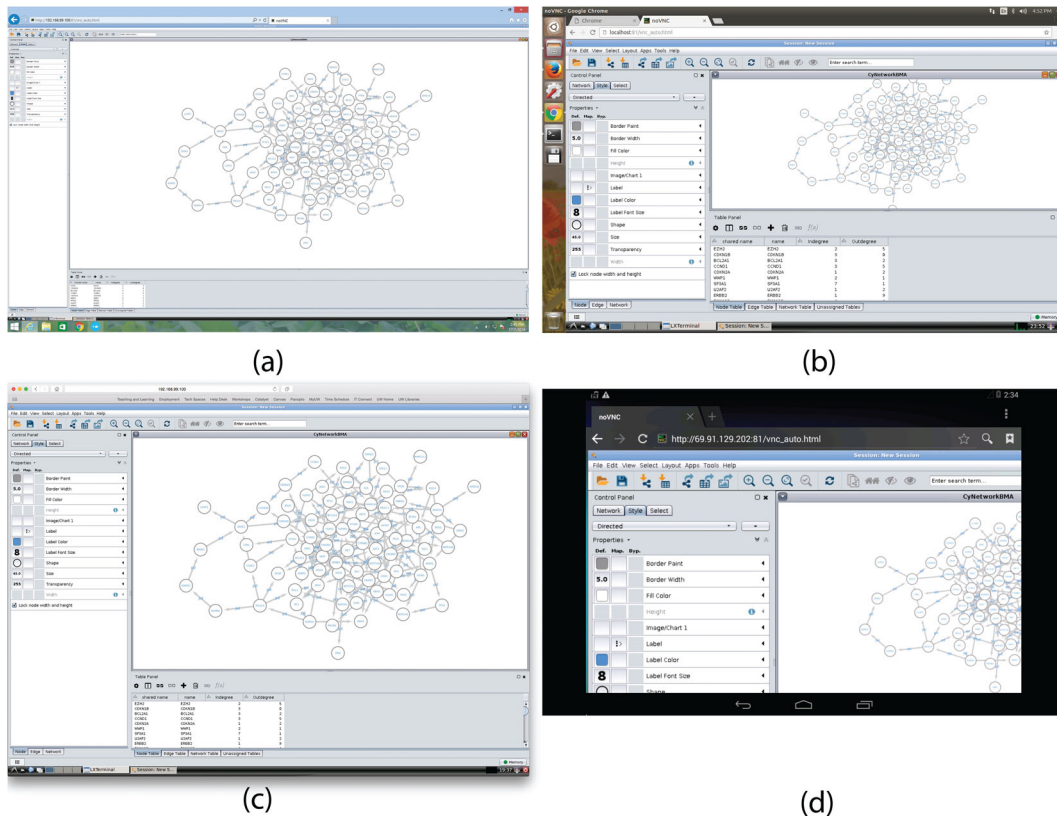
Figure 3: CyNetworkBMA on various browsers and operating systems: **(a)** Internet Explorer on Windows 8.1, **(b)** Google Chrome on Ubuntu Linux, **(c)** Safari on Mac OS, and **(d)** Google Chrome on Android. Given that the interaction to the container is made through a browser, any device with a browser supporting HTML5 will have consistent user experience and results, even on mobile devices such as Android and iOS devices. In the case of mobile devices, the Docker container runs on remote machines (e.g., on a virtual machine instance on cloud).

The tool is extensible and can be used to define all Dockerfile parameters as standard run commands.

Applications

We illustrated the utility of GUIdock-VNC in a proof-of-concept case study of gene network inference. Specifically, we applied GUIdock-VNC to a RNA-seq dataset consisting of 675 human cancer cell lines [19]. We downloaded the variance-stabilized version of the normalized RNA-seq data and extracted a subset of 84 genes that belong to 21 cancer-related pathways (see Supplementary Table 12 in Klijn et al. [19]). We applied the ScanBMA [12] gene network inference algorithm as implemented in the CyNetworkBMA app from within the GUIdock-VNC container.

We show that we get identical results after deploying the package across different browsers on different operating systems. Fig. 3 shows screenshots of using (a) Internet Explorer on Windows 8.1, (b) Google Chrome on Ubuntu Linux, (c) Safari on Mac OS, and (d) Google Chrome on Android. To summarize, we demonstrate the reproducibility of analytical results when GUIdock-VNC is deployed on different browsers and different operating systems.

Benchmarking computational efficiency

Since we have added extra services to the container, it is essential to investigate the performance overhead introduced by these additional services and the container. We conducted an extensive empirical study on comparing performance over different platforms with different hypervisors using GUIdock-X11 and GUI-VNC. As a baseline, we compared the performance of running GUIdock-X11 and GUIdock-VNC on Docker containers to running the CyNetworkBMA app natively. In addition, we also compared our results to running the CyNetworkBMA via a virtual machine (VM). We tested the performance of each of these four options (native, GUIdock-X11, GUIdock-VNC, VM) on the Linux, Macintosh, and Windows operating systems.

In our benchmarking experiments, we used the time series data of the first network from the DREAM 4 crowdsourcing challenge [20,21]. This simulated dataset consists of 100 genes across 21 time points. In order to account for variability in our empirical experiments, we repeated each configuration, i.e., each (OS, option) pair four times. These replicated experiments are represented by "RUN1," "RUN2," "RUN3," "RUN4" in Table 2. In addition, we added warm-up runs to ensure steady-state execution time.

Table 2 shows a consistent minimal overhead of running the proposed container, which is only marginally higher than running the application natively. In particular, we computed the ratio of the average execution time over the four runs to the "native" baseline execution time. Fig. 4 shows the ratio of the average execution time of each of GUIdock-X11, GUIdock-VNC, and VM to the baseline "native" on the Linux, Mac OS, and Windows operating systems. On the Linux, Mac OS, and Windows operating systems, we observed comparable execution time for both GUIdock-X11 and GUIdock-VNC.

Table 2: Execution time in empirical study across the Linux, Macintosh, and Windows operating systems

Platform	Environment	RUN 1	RUN 2	RUN 3	RUN 4	Average	Ratio
Linux	Native	104	100	97	102	100	1
	GUIdock-X11	130	131	131	130	130	1.29
	GUIdock-VNC	134	133	133	134	134	1.39
	VM	187	185	186	185	186	1.92
Mac OS	Native	97	97	96	96	96	1
	GUIdock-X11	120	116	118	123	119	1.23
	GUIdock-VNC	123	121	120	124	122	1.26
	VM	148	143	151	150	148	1.54
Windows	Native	125	127	130	128	127	1
	GUIdock-X11	155	157	155	156	155	1.22
	GUIdock-VNC	157	160	162	161	160	1.25
	VM	179	179	182	184	181	1.42

"Native" means running the CyNetworkBMA app natively on the corresponding OS. "VM" means running the CyNetworkBMA app from a virtual machine on the corresponding OS. The column "average" is the average execution time over the four runs. The column "ratio" is the ratio of the average running time to the "native" baseline.
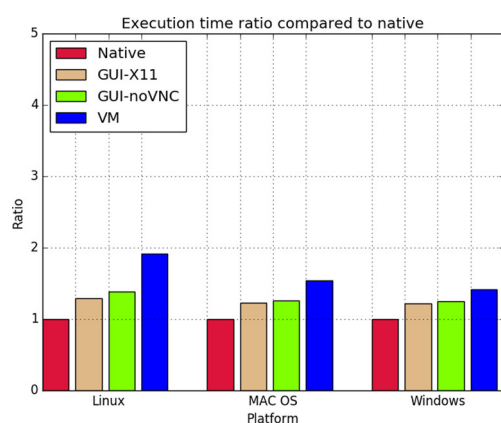
Figure 4: The bar graph shows the ratio of the average execution time to the baseline of running CyNetworkBMA natively for each of Linux, Mac OS, and Windows. The first value for each platform is the ratio of the native execution runtime to itself, and therefore is always equal to 1. The remaining three values correspond to the ratio of the average execution time for GUIdock-X11, GUIdock-VNC, and VM, respectively, to the baseline "native."

Discussion

We present a container tool called GUIdock-VNC that uses a graphical desktop sharing system to provide a browser-based interface for containerized software. The merits of GUdock-VNC are summarized in the following sections.

No installation on client side.

Our proposed container GUIdock-VNC is a self-contained display server with user application and an HTML5-based VNC client, which can be accessed using the web browser. Therefore, any HTML5-capable browser is sufficient on the client side. Almost all modern popular browsers such as Mozilla Firefox, Webkit (Apple Safari and Google Chrome), and Microsoft Edge support HTML5 extensions.

Mobile capable.

Since HTML5 extensions that are required to access noVNC are also supported on mobile versions on modern web browsers, our GUIdock-VNC solution is also accessible through mobile devices such as phones and tablets. The container is a stateful solution, such that it preserves session information even after a disconnected user session. Therefore, researchers can access the container on the go and subsequently continue working on the workstations in the labs.

Cloud integration.

There is a self-contained web server inside the GUIdock-VNC container that is required to access the applications. Our GUIdock-VNC container can be hosted on the cloud by using a reverse proxy or simple NAT rule to pass on any incoming request to the container. We tested the solution with major cloud vendors such as Amazon AWS, Google Cloud, and Microsoft Azure, and the NAT forwarding works without any network interference.

Security using OAuth2.

To ensure secure access to the container, we are using OAuth2 from Google (replaceable by any other identity provider such as LinkedIn, Facebook, Twitter, etc.) to authenticate users by ensuring that the identity registered with the container is owned by the user requesting container access. This can be done by providing email ID, provider user ID and secret password as parameters to the container. If these parameters are provided while initiating the container, the broker will redirect the user to the identity provider to verify the email address. Once authenticated, the identity provider redirects the user to the container.

Availability and requirements

- Project name: GUIdock-VNC
- Project home page: https://github.com/biodepot
- Contents available for download: Docker Images, Dockerfiles, installation scripts, and execution scripts.
- Operating systems: Linux, Mac OS X, Microsoft Windows; specifically, we tested GUIdock-VNC on
  - Linux: Fedora 22/23, Ubuntu 15.04
  - Mac OS X: 10.9, 10.10
  - Microsoft Windows: 7, 8.1, 10
  - Android, IOS
- Programming languages: Python, HTML, JavaScript

- Browsers tested
  – Google Chrome
  – Firefox
  – Safari
  – Microsoft Edge on Windows 10 Pro (using Docker for Windows, and with pop-up blocker off)
- License: MIT License

Availability of supporting data

Snapshots of the code supporting this article are available in the GigaScience GigaDB repository [22].

Additional Files

Additional file 1 — User manual for GUIdock-VNC.
Additional file 2 — Video of GUIdock-VNC demo
   Available on YouTube: https://youtu.be/iaVPnLhOLg0.
Additional file 3 — Video of deploying GUIdock-VNC on the cloud

Abbreviations

GUI, graphical user interface; NAT, network address translation; VM, virtual machine; VNC, Virtual Network Computing.

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

V.M. is the primary developer for GUIdock-VNC. K.Y.Y. coordinated the manuscript preparation. V.M., L.H.H., and K.Y.Y. drafted the manuscript. L.H.H. designed the benchmarking experiments. V.M., L.H.H., D.K., J.K., and S.B.L performed the benchmarking experiments. D.K. and S.B.L. contributed to the comparison of GUIdock-VNC and GUIdock-X11. V.M., J.K., and S.B.L. contributed to the writing of the user manual. V.M. and D.K. made the videos in additional data files. All authors tested GUIdock-VNC and read and approved the final manuscript.

Acknowledgements

References

1. Moreews F, Sallou O, Ménager H et al. Bioshadock: a community driven bioinformatics shared docker-based tools registry. F1000Res 2015;**4**:1443.
2. Bio Docker: Docker for bioinformatics. http://biodocker.org/.
3. Hung LH, Kristiyanto D, Lee SB, Yeung KY. Guidock: using docker containers with a common graphics user interface to address the reproducibility of research. PLoS One 2016;**11**(4):0152686.
4. Richardson T, Stafford-Fraser Q, Wood KR, et al. Virtual network computing. IEEE Internet Computing 1998;**2**(1):33.
5. noVNC by kanaka. http://kanaka.github.io/noVNC/ (9 September 2016, date last accessed).
6. OAuth. http://oauth.net/ (9 September 2016, date last accessed).
7. Docker for Mac and Windows Beta: the simplest way to use Docker on your laptop . https://blog.docker.com/2016/03/docker-for-mac-windows-beta/ (27 December 2016, date last accessed).
8. MobaXterm: enhanced terminal for Windows X11 server tabbed SSH client network tools and much more . http://mobaxterm.mobatek.net/ (9 September 2016, date last accessed).
9. socat: multipurpose relay(SOcket CAT). http://www.dest-unreach.org/socat/doc/socat.html (9 September 2016, date last accessed).
10. Yeung KY, Dombek KM, Lo K et al. Construction of regulatory networks using expression time-series data of a genotyped population. Proc Natl Acad Sci 2011;**108**(48):19436–41.
11. Lo K, Raftery A, Dombek, K et al. Integrating external biological knowledge in the construction of regulatory networks from time-series expression data. BMC Syst Biol 2012;**6**(1):101.
12. Young WC, Raftery AE, Yeung KY. Fast Bayesian inference for gene regulatory networks using ScanBMA. BMC Syst Biol 2014;**8**(1):47.
13. Yeung KY, Fraley C, Young WC, et al. Bayesian model averaging methods and R package for gene network construction. In: Big Data Analytic Technology For Bioinformatics and Health Informatics (KDDBHI), Workshop at the 20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), New York. 2014, https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxr ZGRiaGl8Z3g6NWU1ZGYzZDU4ZDYwNWNhYw.
14. Fronczuk M, Raftery AE, Yeung KY. Cynetworkbma: a cytoscape app for inferring gene regulatory networks. Biology and Medicine 2015;**10**:11 .
15. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;**13**(11):2498–504.
16. Christmas R, Avila-Campillo I, Bolouri H, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. AACR Education Book 2005:12–6. http://educationbook.aacrjournals.org/cgi/content/full/2005/1/12.
17. Cline MS, Smoot M, Cerami E, et al. Integration of biological networks and gene expression data using cytoscape. Nat Protoc 2007;2366–82.
18. Xvfb:virtual framebuffer X server for X version 11 . https://www.x.org/releases/X11R7.6/doc/man/man1/Xvfb.1.xhtml.
19. Klijn C, Durinck S, Stawiski EW, et al. A comprehensive transcriptional portrait of human cancer cell lines. Nat Biotechnol 2015;**33**:306–12.
20. Marbach D, Schaffter T, Mattiussi C, et al. Generating realistic in silico gene networks for performance assessment of reverse engineering methods. J Computat Biol 2009; **16**(2):229–39.
21. Marbach D, Prill RJ, Schaffter T, et al. Revealing strengths and weaknesses of methods for gene network inference. Proc Natl Acad Sci U S A 2010;**107**(14):6286–91.
22. Mittal V, Hung L, Keswani J, et al. 2016. Supporting data for "GUIdock-VNC: using a graphical desktop sharing system to provide a browser-based interface for containerized software." GigaScience Database. http://dx.doi.org/10.5524/100261.