SCIENTIFIC REP🅞RTS

OPEN

NetDiff – Bayesian model selection for differential gene regulatory network inference

Thomas Thorne

Differential networks allow us to better understand the changes in cellular processes that are exhibited in conditions of interest, identifying variations in gene regulation or protein interaction between, for example, cases and controls, or in response to external stimuli. Here we present a novel methodology for the inference of differential gene regulatory networks from gene expression microarray data. Specifically we apply a Bayesian model selection approach to compare models of conserved and varying network structure, and use Gaussian graphical models to represent the network structures. We apply a variational inference approach to the learning of Gaussian graphical models of gene regulatory networks, that enables us to perform Bayesian model selection that is significantly more computationally efficient than Markov Chain Monte Carlo approaches. Our method is demonstrated to be more robust than independent analysis of data from multiple conditions when applied to synthetic network data, generating fewer false positive predictions of differential edges. We demonstrate the utility of our approach on real world gene expression microarray data by applying it to existing data from amyotrophic lateral sclerosis cases with and without mutations in C9orf72, and controls, where we are able to identify differential network interactions for further investigation.

There has been much work in the literature on the inference of networks from gene expression data, utilising a variety of approaches including tests for correlation[1], graphical models[2,3], and mutual information based approaches[4]. Graphical modelling approaches have the advantage that they consider conditional independence, so that regulatory network interactions are only predicted where the other explanatory variables are unable to account for all of the variation observed.

A task that has received less attention, but would allow for greater exploitation of the experimental data that is often generated, is the inference of differential networks[5–9]. The aim of differential network inference approaches is similar to that of tests for differential expression, but rather than detecting changes in the expression of a single gene, the aim is to infer changes in the regulatory network structure itself between multiple conditions. This allows for more specific mechanistic insights into the molecular processes underlying changes between, for example, cases and controls in a disease study[6,8].

In previous work we have explored the application of flexible non-parametric models of regulatory networks that allow for changes in network structure with an arbitrary number of networks[8]. However this flexibility adds complexity to the model that does not necessarily deliver greater insights, at the expense of a complex and computationally demanding implementation. In a Bayesian framework, a parsimonious approach would seem to be applying Bayesian model selection to compare a model allowing for a change in the network structure between conditions against a model of a static network. Bayesian model selection approaches automatically compensate for changes in the number of parameters between models, so that a more complex model will not always improve on a simpler nested model by overfitting the data.

Although differential networks can be learnt through the independent analysis of samples of gene expression data from multiple conditions, as the process suffers from the "*large p small n*" problem, it would be expected that a substantial proportion of the differential edges observed would be a side-effect of false positive or false negative predictions. Instead we propose to compare a model of no change between conditions, against a model where edges are allowed to vary, for each gene independently. This allows the network edges of genes whose interactions do not vary to be learnt more robustly by utilising the data from both conditions, whilst where there is evidence of a change in the network, differential edges are identified.

Division of Brain Sciences, Imperial College London, UK. Correspondence and requests for materials should be addressed to T.T. (email: t.thorne@imperial.ac.uk)

Here we propose a fast and easily applicable method to perform Bayesian model selection on gene expression microarray data collected between multiple conditions, and demonstrate an application of this method to publicly available gene expression microarray data from Amyotrophic Lateral Sclerosis (ALS) cases and controls, investigating the impact of C9orf72 GGGGCC repeat expansions in the underlying regulatory network by comparing patients with and without these mutations.

Methods

Gaussian graphical models.   In a Gaussian Graphical Model (GGM) a set of random variables $X^1, \ldots, X^p$ are assumed to take a multivariate normal distribution so that $X \sim \mathcal{N}(\mu, \Sigma)$. Then in a graphical representation of such a model, pairs of nodes i and j have no edge between them if and only if the two corresponding random variables are conditionally independent, or equivalently $\Sigma^{-1}_{ij} = 0$. Identifying such relationships is possible using the partial correlation between pairs of random variables X^i and X^j with respect to the remaining variables $Z^{-i,j}$, defined as

$$\rho_{X^i X^j \cdot Z^{-i,j}} = \frac{-\Sigma^{-1}_{i,j}}{\sqrt{\Sigma^{-1}_{i,i}\Sigma^{-1}_{j,j}}},$$

(1)

so that an edge from variable i to j exists in the graphical model iff $\rho_{X^i X^j \cdot Z^{-i,j}} \neq 0$. Unfortunately in many cases of interest, such as the analysis of gene expression microarray data, we are faced with the so called "*large p small n*" problem, characterised by a large number of variables p and a comparatively small number of observations n. In such cases it is not possible to directly calculate the partial correlations between the variables as the empirical covariance matrix will be singular. Fortunately however there are a number of ways that we can determine the partial correlations in such cases[10].

One such method to infer the structure of a GGM when we are faced with more variables than observations is to use regularised regression methods to infer the partial correlations between the nodes. If the vector of n observations of variable i is denoted $Y = X^i$ and the remaining variables as $X^{-i} = X^{j \neq i}$, then we seek to infer the posterior of β^i with

$$Y = X^{-i}\beta^i + \varepsilon,$$

(2)

where $\varepsilon \sim \mathcal{N}(0, \sigma^2)$. Then given β^i for each variable i, we can estimate the partial correlations between variables i and j as ref. 10

$$\rho_{ij \cdot Z^{-ij}} = \begin{cases} \text{sign}(\beta^i_j)\ \min\left(1, \sqrt{\beta^i_j \beta^j_i}\right) & \text{if } \text{sign}(\beta^i_j) = \text{sign}(\beta^j_i), \\ 0 & \text{otherwise.} \end{cases}$$

(3)

Here we apply the regularised regression method introduced in Caron *et al.*[11], which has been demonstrated to outperform the lasso and other regularised regression methods. Sparsity is enforced by introducing a scale mixture of Gaussian distributions as the prior on the regression coefficients β_w[11] (omitting indices of β, ζ^2 and σ^2 over the m variables for clarity),

$$p(\beta_w) = \int \mathcal{N}(\beta_w; 0, \zeta^2_w)p(\zeta^2_w)d\zeta^2_w,$$

(4)

with each ζ^2_w in turn distributed as an inverse Gaussian,

$$\zeta^2_w \sim \mathcal{IG}(\alpha, \gamma),$$

(5)

where the parameters α and γ control the penalty placed on the β_w, here set to 0.1 and 1 respectively. We place a vague gamma prior on σ^2. To evaluate equation 3 we take the maximum a posteriori (MAP) estimates of β^i_j to derive partial correlations ρ. Significant partial correlations are identified using the `fdrtool` R package[12] to estimate the local false discovery rate, with a cutoff of 0.025.

Differential networks.   To test for variation in the network structure between conditions, we allow the n observations to be partitioned into K sets, $X_1, \ldots, X_K$. Here we focus on the case where $K = 1$ or $K = 2$, being models of a single network for all observations, and a network varying between two conditions, respectively. Considering the regression approach described in the previous section it is straightforward to allow for independent network structures between the K sets of observations,

$$p(\beta_k, \zeta^2_k, \sigma^2 | Y_k, X_k) \propto \prod_k^K \left[p(Y_k | X_k, \beta_k, \sigma^2) \prod_w^W (p(\beta_{kw}|\zeta^2_{kw})p(\zeta^2_{kw})) \right] p(\sigma^2),$$

(6)

and for a given model M we can then consider the model evidence,

$$p(X|M) = \int_\Theta p(X|\Theta, M)p(\Theta|M)d\Theta,$$

(7)

where the model parameters are subsumed into Θ. Comparing the two competing models $K = 1$ and $K = 2$, we can calculate a Bayes factor $\frac{p(X|M_{K=2})}{p(X|M_{K=1})}$ to assess if there is significant evidence for preferring one model over the

other, and in our method we take a Bayes factor larger than 2 as support for model $M_{K=2}$. Having calculated Bayes factors for each variable $X_1, \ldots, X_p$ we can derive partial correlations for the two conditions $k=1$ and $k=2$ from the appropriate β_{kw}.

Data: Gene expression microarray data X for M genes, a partition of X into two subsets, number of iterations to run variational inference S.

for $i \in 1 \ldots M$ **do**

    Partition observations $Y = X^i$ and X^{-i} into Y_1, X_1^{-i} and Y_2, X_2^{-i};

    Initialise variational parameters for the two partitions, β_1, ζ_1^2, β_2, ζ_2^2 and σ^2;

    for $s \in 1 \ldots S$ **do**

        Update variational parameters for β_1 based on ζ_1^2, σ^2, Y_1, X_1^{-i};

        Update variational parameters for ζ_1^2 based on β_1, σ^2, Y_1, X_1^{-i};

        Update variational parameters for β_2 based on ζ_2^2, σ^2, Y_2, X_2^{-i};

        Update variational parameters for ζ_2^2 based on β_2, σ^2, Y_2, X_2^{-i};

        Update variational parameters for σ^2 based on β, ζ^2, Y, X;

    end

    Set $A^1 \leftarrow \text{MAP}(\beta_1)$;

    Set $A^2 \leftarrow \text{MAP}(\beta_2)$;

    Calculate model evidence $p(X|M_{K=2})$;

    Initialise variational parameters for β, ζ^2 and σ^2;

    for $s \in 1 \ldots S$ **do**

        Update variational parameters for β based on ζ^2, σ^2, Y, X^{-i};

        Update variational parameters for ζ^2 based on β, σ^2, Y, X^{-i};

        Update variational parameters for σ^2 based on β, ζ^2, Y, X^{-i};

    end

    Set $B \leftarrow \text{MAP}(\beta)$;

    Calculate model evidence $p(X|M_{K=1})$;

    if $\frac{p(X \,|\, M_{K=2})}{p(X \,|\, M_{K=1})} > 2$ **then**

        Set $C_i^1 \leftarrow A^1$;

        Set $C_i^2 \leftarrow A^2$;

    else

        Set $C_i^1 \leftarrow B$;

        Set $C_i^2 \leftarrow B$;

    end

end

for $i \in 1 \ldots M$ **do**

    for $j \in 1 \ldots M$ **do**

        Calculate $\rho^1_{ij \cdot Z^{-ij}}$ from C^1;

        Calculate $\rho^2_{ij \cdot Z^{-ij}}$ from C^2;

    end

end

Apply `fdrtool` to estimate local FDR of all $\rho_{ij \cdot Z^{-ij}}$;

Generate network for condition 1 with M nodes, and edges (i, j) where local FDR of $\rho^1_{ij \cdot Z^{-ij}} < 0.025$;

Generate network for condition 2 with M nodes, and edges (i, j) where local FDR of $\rho^2_{ij \cdot Z^{-ij}} < 0.025$;

Algorithm 1: Outline of the inference procedure. Variational parameter updates and model evidence are given in the Supplementary methods. MAP refers to the Maximum A Posteriori value.

Implementation.   To allow us to derive approximate posterior distributions of the model parameters, and perform model selection by estimating the model evidence, we apply a variational Bayes[13–17] inference scheme to our model. An outline of the methodology is given in Algorithm 1, and derivations of the variational updates given in the Supplementary methods.

A benefit of the variational inference approach is that as well as providing a fast approximation of the posterior distribution, it allows for the evaluation of a lower bound on the log evidence $p(x|M)$ for a model M, used as an approximation to $p(x|M)$ in calculating the Bayes factor. The model evidence $p(x|M)$ is not typically available when utilising Markov Chain Monte Carlo (MCMC) methods, and so model selection using MCMC methods typically requires more complex approaches such as Reversible Jump MCMC[18].
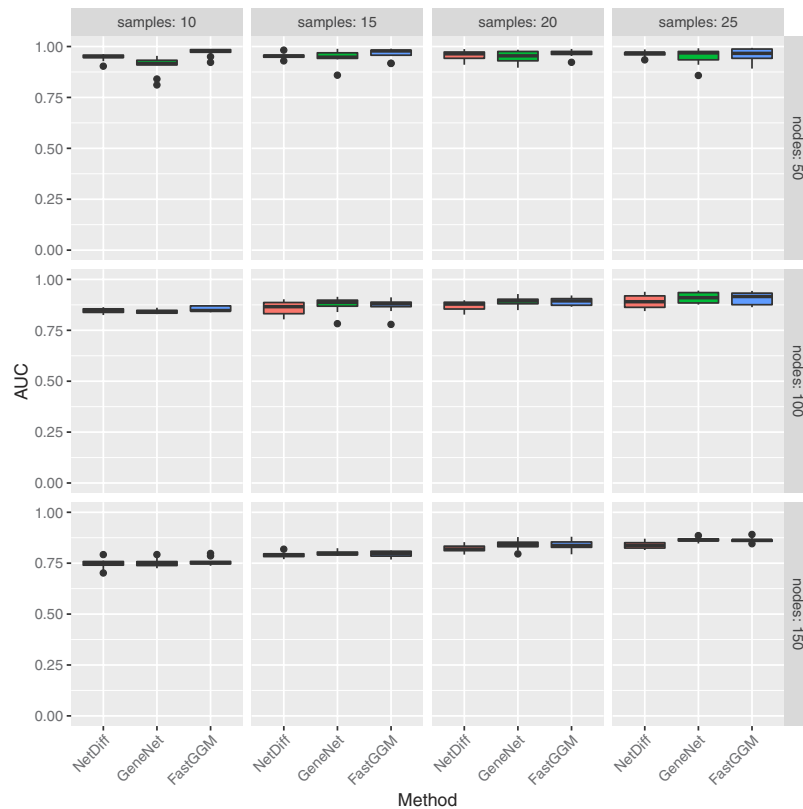
Figure 1. Boxplots of Area Under Curve (AUC) of the receiver operating characteristic curve for each of the methods compared on random networks, with various numbers of network nodes and simulated sample data points from each network.

The method described was implemented as a package for the R Statistical Computing Environment[19] and is available from https://gitlab.com/tt104/NetDiff. In benchmarking the method we also apply the GeneNet[2] and FastGGM[3] R packages.

Synthetic network benchmarking. The `GeneNet`[2] R package was used to simulate partial correlations between sets of variables of varying sizes, and to generate observations of the variables under these network structures. To assess performance in network inference we calculate the area under the Receiver Operating Characteristic (ROC) curve using the `PRROC` R package[20]. This was repeated for multiple synthetic network structures to generate the distributions of metrics shown in the boxplots in Figs 1, 2 and 3.

Results
Synthetic data. We compare our approach to two existing methods from the literature to learn Gaussian graphical models from gene expression microarray data, GeneNet and FastGGM. The running time of our method, and of GeneNet and FastGGM was benchmarked on the synthetic data, see Supplementary Fig. 1. Although our method is slower, the running times for all methods are short, on the order of seconds, so that in practice this will be unlikely to have any real impact. The variational inference approach is considerably more computationally efficient when compared to MCMC approaches, for example in a previous method we developed applying a similar sparse regression method in a Gibbs sampling framework, inference took around 20 hours running in parallel on a 12 core CPU[8].

Network inference. Applying our method to synthetic data with varying network sizes and numbers of observations, we see from the plots in Fig. 1 that our method performs well in terms of the area under the ROC curve, and is comparable to the other approaches that were benchmarked. Although the performance of our method does not exceed that of competing methods, by working in a Bayesian framework we are easily able to expand the approach to perform Bayesian model selection, as described in the following section. There are further benefits to the Bayesian approach, such as the straightforward inclusion of prior information. As expected, the performance of the various approaches decreases as the number of samples is reduced, and also falls as the size of the network grows for a given number of samples.

Differential networks. We test the main focus of our method, the learning of differential networks though Bayesian model selection, by generating synthetic data sets where the network structure is known to change for
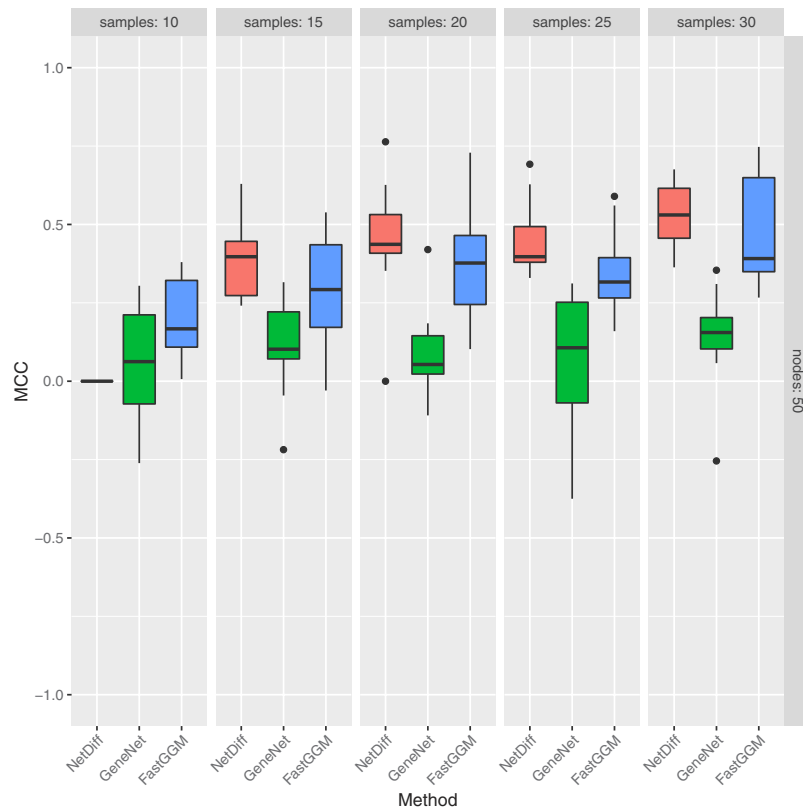
Figure 2. Boxplots of Matthews Correlation Coefficient (MCC) for each of the methods applied to random networks with structures varying between conditions, with various different numbers of samples per condition.

a subset of nodes between two sets of samples, in networks of 50 nodes. To benchmark our method we calculate the Matthews Correlation Coefficient (MCC), comparing the nodes with predicted changes to the known set of positives. Matthews correlation coefficient allows us to test the performance of our method at a specific cut-off, without concerns about effects of class imbalances. To test our hypothesis that a method performing integrated analysis between conditions and attempting to identify evidence for a change in network structure will perform better than independent analyses of each condition, we also compare the performance of GeneNet and FastGGM in identifying nodes whose network interaction partners change between conditions, taking significant edges as those with a local false discovery rate under 0.025 and testing for changes in the inferred networks between conditions.

From the results shown in Fig. 2, it is apparent that for 15 or more samples our approach outperforms the alternative methods, demonstrating higher Matthews correlation coefficient values than GeneNet, and to a lesser extent FastGGM. Our approach has less variance in performance than FastGGM, likely due to the greater susceptibility to noise inherent in learning differential networks from independent analyses. However our method does require a relatively large number of samples per condition to effectively predict changes in the network structure, as can be seen in the case for 10 samples.

To further test the robustness of our approach, we also apply the method to synthetic data from two conditions, where the network structure does not change between conditions, as shown in Fig. 3. As there are no true positive differential edges, we only measure the false positive rate, which for our method is significantly lower than applying GeneNet or FastGGM to each condition independently. This is to be expected, as our method is able to identify that there is no change in the network structure, and learn using data from both conditions combined, whilst the other approaches must learn networks independently from each condition, and differential edges may be erroneously identified simply due to noise in the output.

Dysregulation of gene expression in ALS patients with C9orf72 mutations. To test the applicability of our approach to real world datasets we applied our method to the data of Cooper-Knock et al.[21], Gene Omnibus accession GSE68607, consisting of gene expression microarray data collected from lymphoblastoid cell lines derived from 54 ALS patients and 15 controls. The ALS patients are further divided into 31 with C9orf72 mutations and 23 without. Hexanucleotide repeat expansions in C9orf72 have been shown to cause familial ALS[22], but the exact mechanisms are not understood. To identify the causal mechanisms by which these expansions lead to ALS, we seek to investigate the differences in gene regulation between controls and ALS cases with and without expansions. We do so by comparing gene expression microarray data from each subclass of ALS patients to one another, and to controls, identifying differential network edges using our proposed method. To
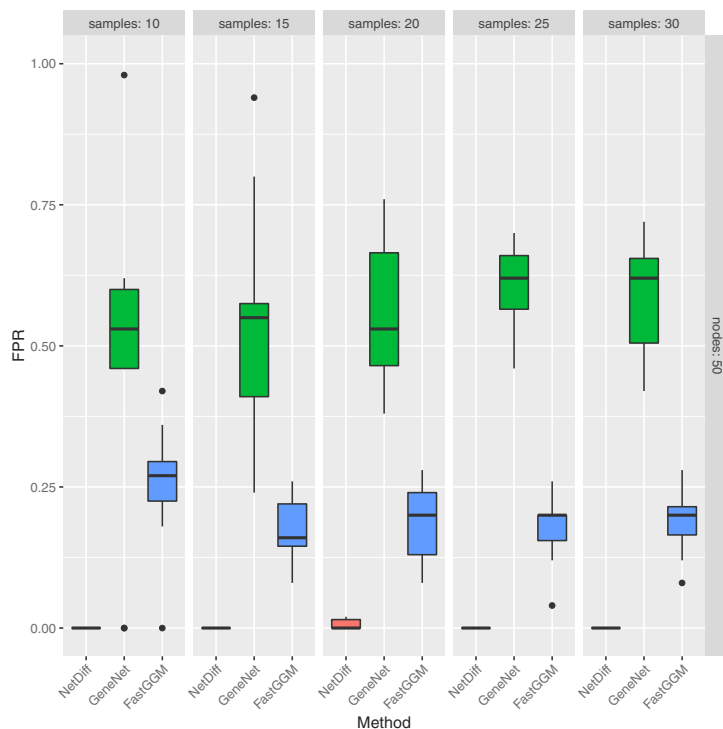
Figure 3. Boxplots of False Positive Rate (FPR) for each of the methods applied to random networks with structures remaining identical between conditions, with various different numbers of samples per condition.
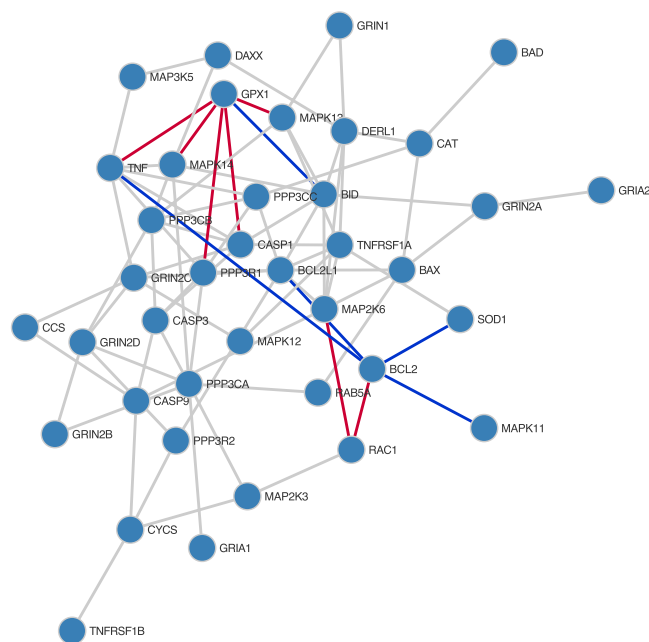

Figure 4. Differential network comparing ALS cases without C9orf72 mutations and controls. Conserved network edges are shown in grey, whilst edges only present in ALS cases are in red, and those only present in controls are in blue.

render the analysis tractable we choose a subset of the genes in the dataset to focus on in our study, specifically those in the KEGG database[23,24] ALS pathway, accession *h*sa05014, as well as C9orf72.

Firstly considering the differential networks inferred between cases without C9orf72 mutations and controls, as shown in Fig. 4, we see that whilst the majority of network edges are conserved, there appear to be a small
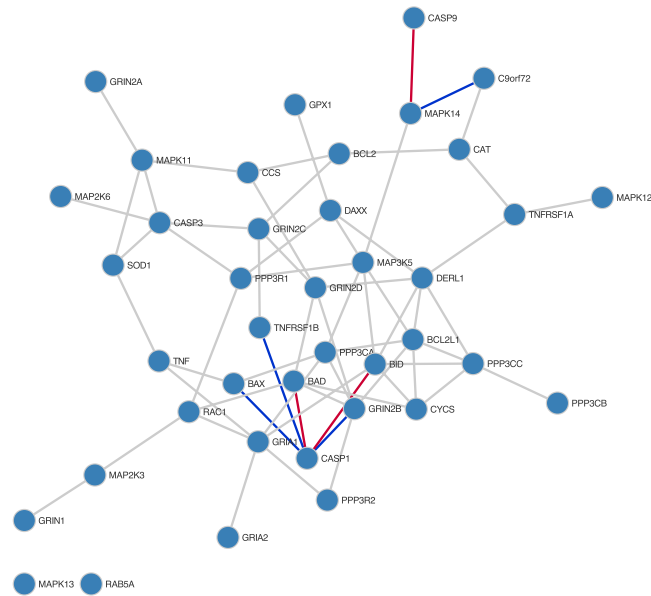
Figure 5. Differential network comparing ALS cases with and without C9orf72 mutations. Conserved network edges are shown in grey, whilst edges only present in ALS cases with C9orf72 mutations are in blue, and those only present in ALS cases without C9orf72 mutations are in red.
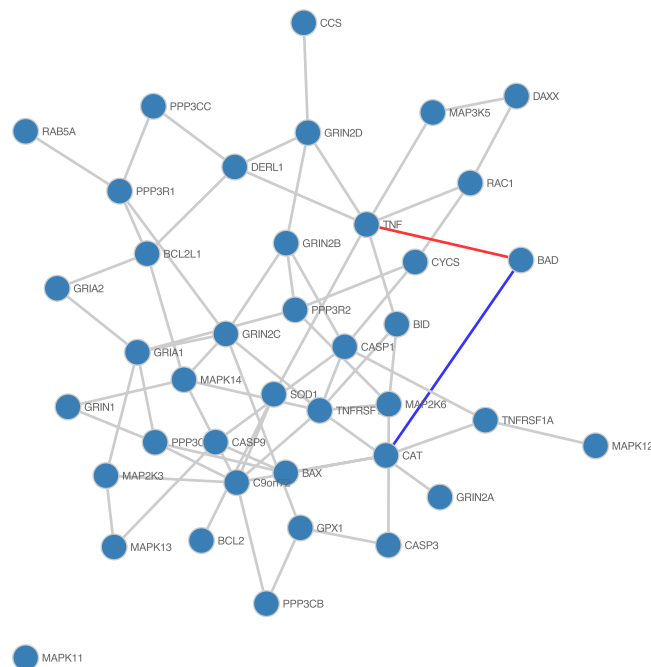

Figure 6. Differential network comparing ALS cases with C9orf72 mutations and controls. Conserved network edges are shown in grey, whilst edges only present in ALS cases with C9orf72 mutations are in red, and those only present in controls are in blue.

number of genes whose interactions change considerably. Several regulatory interactions are gained by GPX1, to PPP3R1, CASP1, TNF, all promoters of apoptosis, and MAPK13 and MAPK14. The interaction between BID and GPX1 is only observed in controls. BCL2 loses interactions with SOD1, BCL2L1, TNF and MAPK11, whilst gaining an interaction with RAC1. Interestingly, BCL2 is crucial in inhibiting apoptosis, and has been demonstrated to have reduced expression in ALS in mice[25]. RAC1, thought to be a key player in immune response in ALS[26], also gains an interaction with MAP2K6.

Changes within the set of ALS patients, between those with and without mutations in C9orf72, are less pronounced, as seen in Fig. 5. Interestingly cases with a C9orf72 mutation gain an interaction with MAPK14, whilst

MAPK14 and CASP9 do not interact in cases with a C9orf72 mutation. Furthermore CASP1 gains interactions with BAX, TNFRSF1B and GRIN2B, whilst losing interactions with BAD and BID. CASP1, BAD, BAX and BID all promote apoptosis, whilst TNFRSF1B inhibits apoptosis. GRIN2B encodes NR2B, which activates apoptosis when overexpressed[27].

We also compare patients with mutations in C9orf72 with controls, again giving less pronounced results, shown in Fig. 6. In patients with C9orf72 mutations, BAD appears to lose an interaction with CAT, and gain one with TNF. As mentioned previously TNF is a promoter of apoptosis, again showing dysregulation of apoptosis as a potential mechanism of ALS[28].

Genes having gene ontology annotations for positive regulation of apoptosis were found to be significantly differentially expressed in microarray data from lymphoblastoid cell lines derived from ALS cases with C9orf72 mutations[21], and our results suggest that there may be differences in regulation of apoptosis between ALS cases with and without C9orf72 mutations.

It should be highlighted that the network edges (interactions) in inferred graphical models from gene expression data do not necessarily correspond to direct regulatory interactions in the form of transcription factor binding, but may take the form of the indirect modulation of transcription through the action of intermediary steps that are not measured by transcriptomic studies.

Discussion

We demonstrate that our Bayesian model selection approach is able to identify changes in network structures between conditions. Whilst the performance of our approach in learning the network structure under a single condition does not improve on competing methods, the focus is in this work to allow for the inference of differential networks to allow gene expression microarray data from case control studies to be fully exploited. The model selection approach we describe here is enabled by the variational inference scheme we employ that allows us to estimate the model evidence in a computationally efficient manner, and we show that it allows us to more robustly identify differential regulatory interactions between conditions. Due to the challenging nature of the task of learning a network from microarray data, likely introducing noise in the results, an integrated approach that can utilise all of the samples whilst testing for changes in the network structure is preferable to independent learning on cases and controls, as we demonstrate here.

References

1. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9,** 1 (2008).
2. Opgen-Rhein, R. & Strimmer, K. From correlation to causation networks: a simple approximate learning algorithm and its application to high-dimensional plant gene expression data. *BMC Systems Biology* **1,** 1, 37 (2007).
3. Wang, T. *et al.* FastGGM: An Efficient Algorithm for the Inference of Gaussian Graphical Model in Biological Networks. *PLOS Computational Biology* **12,** e1004755 (2016).
4. Margolin, A. A. *et al.* ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics* **7** Suppl 1, S7 (2006).
5. Bandyopadhyay, S. *et al.* Rewiring of Genetic Networks in Response to DNA Damage. *Science* **330,** 1385–1389 (2010).
6. West, J., Bianconi, G., Severini, S. & Teschendorff, A. E. Differential network entropy reveals cancer system hallmarks. *Scientific Reports* **2,** 802 (2012).
7. Ideker, T. & Krogan, N. J. Differential network biology. *Molecular Systems Biology* **8** (2012).
8. Thorne, T. *et al.* Graphical modelling of molecular networks underlying sporadic inclusion body myositis. *Mol. BioSyst.* **9,** 1736–1742 (2013).
9. Okawa, S., Angarica, V. E., Lemischka, I., Moore, K. & del Sol, A. A differential network analysis approach for lineage specifier prediction in stem cell subpopulations. *npj Systems Biology and Applications* **1,** 15012 (2015).
10. Krämer, N., Schäfer, J. & Boulesteix, A. L. Regularized estimation of large-scale gene association networks using graphical Gaussian models. *BMC Bioinformatics* **10,** 384 (2009).
11. Caron, F. & Doucet, A. Sparse Bayesian nonparametric regression. In *ICML '08,* 88–95 (ACM, New York, New York, USA, 2008).
12. Strimmer, K. fdrtool: a versatile R package for estimating local and tail area-based false discovery rates. *Bioinformatics* **24,** 1461–1462 (2008).
13. MacKay, D. J. C. Developments in Probabilistic Modelling with Neural Networks — Ensemble Learning. In *Machine Learning,* 191–198 (Springer London, London, 1995).
14. MacKay, D. J. C. *Information Theory, Inference and Learning Algorithms* (Cambridge University Press, 2003).
15. Bishop, C. M. *Pattern Recognition and Machine Learning* (Springer Verlag, 2006).
16. Barber, D. *Bayesian Reasoning and Machine Learning* (Cambridge University Press, 2012).
17. Murphy, K. P. *Machine Learning. A Probabilistic Perspective* (MIT Press, 2012).
18. Green, P. J. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* **82,** 711–732 (1995).
19. R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/ (2016).
20. Grau, J., Grosse, I. & Keilwagen, J. PRROC: computing and visualizing precision-recall and receiver operating characteristic curves in R. *Bioinformatics* **31,** 2595–2597 (2015).
21. Cooper-Knock, J. *et al.* C9ORF72 GGGGCC Expanded Repeats Produce Splicing Dysregulation which Correlates with Disease Severity in Amyotrophic Lateral Sclerosis. *PLOS One* **10,** e0127376 (2015).
22. DeJesus-Hernandez, M. *et al.* Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron* **72,** 245–256 (2011).
23. Kanehisa, M., Goto, S., Furumichi, M., Tanabe, M. & Hirakawa, M. KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Research* **38,** D355–60 (2010).
24. Kanehisa, M., Goto, S., Kawashima, S. & Nakaya, A. The KEGG databases at GenomeNet. *Nucleic Acids Research* **30,** 42–46 (2002).
25. Vukosavic, S., Dubois-Dauphin, M., Romero, N. & Przedborski, S. Bax and Bcl-2 interaction in a transgenic mouse model of familial amyotrophic lateral sclerosis. *Journal of neurochemistry* **73,** 2460–2468 (1999).
26. Dâ Ambrosi, N., Rossi, S., Gerbino, V. & Cozzolino, M. Rac1 at the crossroad of actin dynamics and neuroinflammation in Amyotrophic Lateral Sclerosis. *Frontiers in Cellular Neuroscience* **8,** 1328 (2014).
27. Liu, Z. *et al.* NR2B-Containing NMDA Receptors Expression and Their Relationship to Apoptosis in Hippocampus of Alzheimer's Disease-Like Rats. *Neurochemical Research* **37,** 1420–1427 (2012).
28. Sathasivam, S. & Shaw, P. J. Apoptosis in amyotrophic lateral sclerosis–what is the evidence? *The Lancet Neurology* **4,** 500–509 (2005).

Acknowledgements

Author Contributions

T.T. conceived and implemented the method, and analysed the results. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The author declares no competing financial interests.

How to cite this article: Thorne, T. NetDiff – Bayesian model selection for differential gene regulatory network inference. *Sci. Rep.* **6**, 39224; doi: 10.1038/srep39224 (2016).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.