RESEARCH

Open Access

In silico discovery of blood cell macromolecular associations



Kaare M. Gautvik¹, Daniel Sachse², Alexandra C. Hinton³, Ole K. Olstad², Douglas P. Kiel⁴, Yi-Hsiang Hsu⁴, Tor P. Utheim^{2,5,6,7}, Christine W. Lary³ and Sjur Reppe^{1,2,5*}

Abstract

Background: Physical molecular interactions are the basis of intracellular signalling and gene regulatory networks, and comprehensive, accessible databases are needed for their discovery. Highly correlated transcripts may reflect important functional associations, but identification of such associations from primary data are cumbersome. We have constructed and adapted a user-friendly web application to discover and identify putative macromolecular associations in human peripheral blood based on significant correlations at the transcriptional level.

Methods: The blood transcriptome was characterized by quantification of 17,328 RNA species, including 341 mature microRNAs in 105 clinically well-characterized postmenopausal women. Intercorrelation of detected transcripts signal levels generated a matrix with > 150 million correlations recognizing the human blood RNA interactome. The correlations with calculated adjusted *p*-values were made easily accessible by a novel web application.

Results: We found that significant transcript correlations within the giant matrix reflect experimentally documented interactions involving select ubiquitous blood relevant transcription factors (CREB1, GATA1, and the glucocorticoid receptor (GR, NR3C1)). Their responsive genes recapitulated up to 91% of these as significant correlations, and were replicated in an independent cohort of 1204 individual blood samples from the Framingham Heart Study. Furthermore, experimentally documented mRNAs/miRNA associations were also reproduced in the matrix, and their predicted functional co-expression described. The blood transcript web application is available at http://app.uio.no/med/klinmed/correlation-browser/blood/index.php and works on all commonly used internet browsers.

Conclusions: Using in silico analyses and a novel web application, we found that correlated blood transcripts across 105 postmenopausal women reflected experimentally proven molecular associations. Furthermore, the associations were reproduced in a much larger and more heterogeneous cohort and should therefore be generally representative. The web application lends itself to be a useful hypothesis generating tool for identification of regulatory mechanisms in complex biological data sets.

Keywords: Blood, Transcriptome, Web application, Molecular associations, Microarrays

*Correspondence: reppe.sjur@gmail.com

¹ Unger-Vetlesen Institute, Lovisenberg Diaconal Hospital, Lovisenberggata 17, 0456 Oslo, Norway

Full list of author information is available at the end of the article



Background

Regulation of human gene expression relies on functional macromolecules, including transcription factors (TFs), and micro RNAs (miRNAs). TFs may induce or suppress transcription of their target genes, exerted via distinct binding sites and interaction with other signalling molecules [1]. miRNA's main function is inactivation of mRNAs [2]. Our hypothesis is that highly correlated

© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/ficenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

transcripts in blood and tissue may reflect important functional associations and be a useful tool for hypotheses generation. Signal molecules are often involved in the same pathways and likely to be similarly regulated [3, 4]. We have used the global transcriptome generated from peripheral blood donated by an Oslo cohort of 105 postmenopausal women who were similar with respect to age, ethnicity and health status to generate a large correlation matrix with Pearson r values and modes of significance based on transcription signal values. A web application was developed to explore associations between genes of interest in the dataset.

Coexpressed transcripts tend to reflect co-expressed proteins [5] and we hypothesized that highly correlated transcripts could reflect associations at the protein level. Three ubiquitously expressed TFs known to be functionally important in blood cells were used for testing: cyclic AMP-responsive element (CRE)-binding protein 1 (CREB1), GATA Binding Protein 1 (GATA1) and glucocorticoid receptor (GR). CREB1 is involved in several aspects of hematopoiesis [6-9]. DNA binding and activation of CREB1 depends on its phosphorylation, for example induced by parathyroid hormone (PTH) [10]. Furthermore, expression of CREB1 in peripheral blood mononuclear cells (PBMC) correlates positively with CREB1 expression in the postmortem brain of Alzheimer's patients [11]. Thus, identification of associated transcripts may help to identify novel TFs and their gene targets of functional importance in blood and other tissues. In fact, three datasets of gene expression in immortalized B cells from normal individuals were used to show that correlated transcript levels could be used to predict gene function [12]. GATA1 is essential for erythroid development by regulating the switch of fetal hemoglobin to adult hemoglobin in haemopoietic cells. GATA1 is a multifunctional gene regulating a plethora of genes, and Encyclopedia of DNA Elements (ENCODE) has registered its response element in nearly 10,000 genes.

The glucocorticoid receptor (GR, *NR3C1*) is also expressed in most tissues and cells, modulating activities of genes involved in cell differentiation/development, metabolism, and immune responses [13]. Natural forms of glucocorticoids or analogs like dexamethasone, all acting through GR, are frequently used medical drugs with direct effects on almost all cell types [14]. Even in physiological concentrations, glucocorticoids regulate major aspects of immune cell functions and are powerful immunosuppressants at pharmacological doses [15].

We also verified the representativeness of our dataset by replicating the top 200 associations in an independent cohort from the Framingham study. Furthermore, we tested if experimentally proven miRNA/mRNA associations, also were statistically correlated in our dataset.

Construction and content *Blood donors*

For the postmenopausal blood sampling, Norwegian women (50–86 years, n = 105) representing a cohort with varying bone mineral densities (BMDs) and free of primary diseases known to affect the skeleton, were consecutively recruited as described [16]. Blood was collected in the morning from fasting individuals. Postmenopausal women from the Offspring cohort (women aged 40–92, n = 1204) participating in The Framingham Study [17, 18] were used as replication cohort. The Framingham Heart Study (FHS) is an ongoing prospective community-based study that includes the children of the original cohort and their spouses, who were enrolled into the Offspring Cohort in 1971. At each FHS examination, age, height, weight and extensive questionnaires were obtained according to standardized protocols. For this analysis, we included Offspring participants who attended examination cycle 8 (2005-2008). Gene expression data was collected for (n = 2442) as previously described [18, 19]. These were further filtered on female sex and menopause to achieve the final sample size (n = 1204). Of note, hormone replacement therapy was not included in filtration criteria.

RNA purification and gene expression analysis

RNA from whole blood was isolated according to the PAXgene Blood RNA Kit manufacturer (BD, Franklin Lakes, NJ, USA), including the optional on-column DNase digestion according to manufacturer's instructions. RNA from the Oslo and Framingham cohorts were analysed according to manufacturer's instructions on the Affymetrix Human Gene 1.0 ST GeneChip (Thermo Fisher Scientific, Waltham, MA, USA) which contains ~ 1.4 million probe sets in total. In brief, the Affymetrix Human Exon 1.0 ST Array (Affymetrix, Inc., Santa Clara, CA) was used and gene annotations were obtained from Affymetrix NetAffx Analysis Center (version 31), resulting in ~ 17,000 distinct genes for downstream analysis.

A PCR based method involving LDA cards A and B was used for quantification of ~700 microRNAs and other non-coding RNAs in the Oslo cohort according to manufacturer's instructions (Thermo Fisher Scientific, Waltham, MA, USA).

Calculations, statistics and the web application

Pearson product-moment correlation coefficients (r) were computed between expressions of all genes (> 17,000 probe sets) across 105 women using \log_2 transformed Affymetrix RMA (Robust Multi-array Average) signal values and inversed PCR Ct values and saved in a database along with their corresponding *p*-values. A

web application publicly available at http://app.uio.no/ med/klinmed/correlation-browser/blood/index.php was programmed in order to access the database and flexibly search for correlations of interest as previously described [20]. A screenshot of the web application is displayed in Fig. 1. Search results are returned together with raw and Bonferroni-corrected *p*-values and a measure of the false discovery rate (FDR) as estimated by the Benjamini & Hochberg procedure [21]. This procedure has been shown to control FDR when the tests are independent or positively correlated [22]. This assumption is reasonable when identifying differentially expressed genes. The Oslo cohort generated expression data earlier and was the basis for development of the web application. When other expression data became available later, the Framingham data was selected for replication because of similarity with regard to the platform. In the replication analysis, we computed Pearson product moment correlations, followed by Bonferroni correction. Algorithm and methods used in generation of the web application have been more thoroughly described in a previous similar paper [20].

Utility

Verifications

First, we identified the 200 most significantly correlated transcripts for each of the selected transcription factors (CREB1, GATA1, GR) and second, tested if associations identified by the web application reflected experimentally verified interactions. Then, we analysed if the genes harboured the corresponding transcription factor binding sites in their promoters. For this we used ENCODE summarizing results from Chromatin Immunoprecipitation (ChIP) studies by use of the Harmonizome web portal (https://maayanlab.cloud/Harmo nizome/gene_set/NR3C1/ENCODE+Transcription+ Factor+Targets) [23].

Promoter elements binding CREB1 protein, were identified in 182 of the 200 (91%) topmost CREB1 correlated genes (Table S1), In contrast, when selecting random genes from the same Oslo cohort, 48.3% (SD=2.1%) were identified as having a CREB1 promoter element (not shown). In all, 13,251 genes have this element as registered by ENCODE, thus, a high fraction of associations are expected for sets of random genes. Similar results were obtained for GATA1 with binding elements found in 73.0% while 34.5% (SD=6.1) binding elements were found in 200 random genes (Table S2). For GR 35.5% of 200 topmost correlated genes had the GR binding element while random genes had 14.0% (SD=3.0) binding sites (Table S3). ENCODE have registered 9608 and 4104 genes with binding sites for GATA1 and GR, respectively.

We tested representativeness of the Oslo cohort by checking if the 200 most significant associations in the Oslo cohort were reproduced in blood from postmenopausal women at exam 8 in Framingham Study (N=1204). For CREB1, 180 associations reached significance also in the Framingham cohort (six did not reach significance, and 14 transcripts were undetected) (Table S1). For GATA1 and GR, all, but five and four of the transcripts detected in both cohorts reached statistical significance in the Framingham cohort (Tables S2 and S3, respectively).

Evaluation of associations involving microRNAs

To verify whether the web application/matrix can identify also putative miRNA targets, we took advantage of experimentally proven miRNA targets in TarBase 8.0 using DIANA Tools (https://carolina.imis.athena-innov ation.gr/diana_tools/web/index.php?r=tarbasev8% 2Findex) [24]. For each of the ten miRNAs most highly expressed in peripheral blood based on their PCR Ctvalues, the 20 best experimentally verified interacting mRNAs, accessing all cell lines and tissues, were obtained. Out of the ~200 interactions/associations, 50 (25%) appeared as nominally significant when analysed in the web application (Table S4). As an alternative evaluation strategy, we selected the top 50 experimentally verified miRNA/mRNA interactions in blood from TarBase 8.0, and found that 30 pairs reached detection level in the Oslo cohort and 13 (43%) of these obtained nominal significance (Table S5).

Analysis of the blood interactome employing ingenuity pathway analysis (IPA)

Initially, we tested whether experimentally proven, functional associations mapped by IPA were reproduced in our data. As presented in Table S6, transcripts associating with CREB1 as well as GATA1 were statistically over-represented in functions related to haematological systems within the category "Physiological System Development and Function" As expected, more

Fig. 1 Interface of the web application. A typical search starts with inserting an identifier in the first window of "Search Options", either Entrez Gene ID (e.g., "1234"), an Accession Number (e.g., "BE644809" or "NM_005715"), a Gene Symbol (e.g., "NR3C1" or "CREB1", not case-sensitive) or an Affymetrix probe set ID (e.g., "8,114,814"). Then, a specific transcript can be traced by inserting a second identifier in window two under "Search Options". Alternatively, the window may be left open to obtain a list of the transcripts most significantly correlating to the identifier in the first window. Filling in boxes in the "Output Options" fields enables restriction of output to e.g., transcripts having specific keywords in Gene Ontology (GO), TFs (genes having "transcription" as part of the Gene Title) or only positive or negative correlations

⁽See figure on next page.)

COL	walati.	ana in	the blee	d data	+				
	relatio	ons in	the bloo		set				
A typica "1234") case-se	al search star), an Accessio ensitive) or ar	ts with insert on Number (e n Affymetrix p	ting an identifier in e.g., "BE644809″ or probe set ID (e.g., '	the first window "NM_005715") "8114814").	w of "Sea), a Gene	arch Options" Symbol (e.g	", either Entrez Ger g., "NR3C1" or "CR	ne ID (e.g., EB1", not	
Then, a Alterna identific	a specific tran tively, the wir er in the first	script can be ndow may be window.	traced by inserting left open to obtain	g a second ident a list of the tra	tifier in w anscripts	vindow two u most signific	nder "Search Optic cantly correlating t	ons". o the	
Filling in in Gene correlat	n boxes in the e Ontology (G tions.	e ``output opt O), TFs (gen	ions″ fields enables es having "transcrip	s restriction of o ption" as part o	output to f the Ger	e.g., transcr ne Title) or o	ripts having specifi nly positive or neg	c keywords ative	
(View t	he full list of	available key	terms [here]).						
		S	earch Options: Fir	nd correlation	s betwe	en			
Gene/l	Label/Number	r: NR3C1	and	all others		(Empty a	also means "all oth	ers")	
			Outp	put Options					
Res	strict output t	o 🗆 miRs,	Transcription fac	ctors? (Empty:	No restri	ctions.)			
		GO: all On	tologies 🗸	with keyword:	anything				
		KEGG Path	nway: any						
,	Limit output t	o 50 i	rows.						
	Remov	e 🗹 "self co	orrelations", i.e. cor	rrelations of a v	ariable v	vith itself			
1	Sort results b	y O positiv	e, \bigcirc negative or	absolute val	ue of cor	relation			
						Clear	form Find correlati	ons	
Starting	g timer.								
Fetching Getting 1 of 1 of Calculat	g correlation list of correla list of p-valu done. ting correctio	data, please ations for id 148 ns for multip	176 NR3C1 wait 14821. 21. le testing (n=1732)	 nuclear 7), please wait, 	r recepto	r subfamily 3	3, group C, membe	r 1 (glucocorticoid	d receptor)
Fetchin Getting Getting I of 1 c Calculai Done; V Colu R L L L L L L L L S R R R R R R R R R R	g correlation list of correlation list of correlation done. ting correction whole program mmn legge tank: Row nur abel1: Affyrm abel2: Affyrm abel2: Affyrm cccNum: The ti seneSym: Offi nfo: Link to m rearson r: Pea taw p: Unadju	data, please ations for id : es for id 148 ns for multip m took 0.793 end: mber sorted i terix probe se unique identi icial gene syn ore informat rison correlat sted p-value	176 NR3C1 wait L4821. 21. le testing (n=1732' seconds. Transmiti by strongest correla at identifier(s) of th at identifiers of gen fier for a sequence mol ion about the varia ion coefficient	 nuclear 7), please wait. 7)	the sear o the Lat ble at NG	r subfamily 3 ch bel 1 gene .BE]	3, group C, membe	r 1 (glucocorticoic	d receptor)
Fetchin Getting Getting I of 1 c Calcular Done; V Colu R LL L L A A G G I Ir P R B F F	I oritidat g correlation l list of correlation done. ting correctio whole program Imm legge tank: Row nur abel1: Affyrme abel2: Affyrme abel2: Affyrme ccNum: The I beneSym: Off beneSym:	data, please ations for id : es for id 148 es for id 148 ns for multip m took 0.793 end: mber sorted i atrix probe si atrix probe si etrix pr	176 NR3C1 wait (4821. 21. le testing (n=1732: seconds. Transmitt by strongest correls et identifier(s) of th at identifiers of gen hol fer for a sequence mol ion about the varia ion coefficient ected p-value rate calculated by t	nuclear 7), please wait. ting 50 results: ation e gene used in es correlating t record [accessible in the giver the Benjamini-F	the sear o the Lat ble at NG n row	r subfamily 3 ch sel 1 gene :BI] Procedure	3, group C, membe	r 1 (glucocorticoic	d receptor)
Fetchin Getting Getting Calcular Done; V Colu R L L L L L L L L G G T F R R R R R R R R R	I Griffer I g correlation I list of correlation I list of p-valu done. ting correctio whole program Imm legge tank: Row nur abel2: Affyrm tabel2:	data, please ations for id : es for id 148 en for multip m took 0.793 end: mber sorted i etrix probe sa unique identi icial gene syr unique identi icial gene syr unique identi icial gene syr unique identi icial gene syr unique identi icial syr end: mber sorted i etrix probe sa unique identi icial gene syr unique identi icial gene syr	176 NR3C1 wait 14821. 21. le testing (n=1732) seconds. Transmiti by strongest correla at identifier(s) of the st identifier(s) of the fier for a sequence nbol ion about the varia ion coefficient ected p-value rate calculated by t AccNum	nuclear 7), please wait, ting 50 results: ation e gene used in es correlating t record [accessible in the giver the Benjamini-F GeneSym	the sear o the Lat ble at NG o row Hochberg Info	r subfamily 3 ch sel 1 gene :BI] Procedure Pearson r	3, group C, membe	r 1 (glucocorticoic	fDR (BH)
Fetchin Getting Getting Done; V Colula Done; V Colul R LL LL LL LL LL B R B F F R R R R R R R R R R R R R R R	I orrelation I list of correlation I list of correlation done. Iting correctio whole program Imm legge tank: Row nur abel2: Affyrm abel2: Affyrm abel2: Affyrm cocNum: The i abel2: Affyrm too Link to m rearson r: Pea aw p: Unadju bonferroni: Bo DR (BH): Fals Label1 8114814	data, please ations for id : es for id 148 ns for multip m took 0.793 end: mber sorted i etrix probe se unique identi icial gene syr unique identi icial gene syr unique identi icial gene syr unique identi icial gene syr unique identi icial sene syn unique identi icial sene syn	176 NR3C1 wait L4821. 21. le testing (n=1732: seconds. Transmiti by strongest correla at identifier(s) of th at identifiers of gen fier for a sequence mbol ion about the varia ion coefficient ected p-value rate calculated by t <u>AccNum</u> NM_018695	() nuclear 7), please wait. ting 50 results: ation te gene used in es correlating t record [access] bble in the giver the Benjamini-P GeneSym ERBB2IP	the sear o the Lat ble at NG n row Hochberg	ch bel 1 gene EBI] Procedure Pearson r 0.924976	Raw p 4.449643e-45	Bonferroni 7.709896e-41	FDR (BH) 7.709896e-41
Fetchin Getting Getting Done; V Coluu R L L L L L L L L L B F F R ank 1 2	I oritidat g correlation l list of correlation done. ting correctio whole program Imm legge tank: Row nur abel: Affyrm abel2: Affyrm uccNum: The i abeneSym: Off baneSym: Off fo: Link to m earson r: Pea aw p: Unadju bonferroni: Bo DR (BH): Fals Label1 S114814 S114814	data, please ations for id : es for id 148 ns for multip m took 0.793 end: mber sorted i etrix probe se unique identi icial gene syr unique identi isted p-value state isted p-value state	176 NR3C1 wait L4821. 21. le testing (n=1732: seconds. Transmiti by strongest correla at identifier(s) of th at identifiers of gen fier for a sequence mbol ion about the varia ion coefficient ected p-value rate calculated by t <u>AccNum</u> NM_018695 NM_018183	() nuclear 7), please wait. ting 50 results: ation te gene used in es correlating t record [access] ble in the giver the Benjamini-F <u>GeneSym</u> ERBB2IP SBN01	the sear o the Lab ble at NG n row Hochberg Info 3	ch bel 1 gene EBI] Procedure Pearson r 0.924976 0.922925	Raw p 4.449643e-45 1.693471e-44	Bonferroni 7.709896e-41 2.934277e-40	FDR (BH) 7.709896e-41 1.467138e-40
Fetchin Getting Calcula Done; V Colu R R R R R R R R R R R R R R R R R R R	I STITEST I g correlation I list of correlation I list of correlation whole program IIII of provention IIII of correction whole program IIIII of correction whole program IIIII of correction IIIII of correction IIIII of correction IIIII of correction IIIII of correction IIIIII of correction IIIIII of correction IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	data, please ations for id 1 es for id 148 ns for multip m took 0.793 end: mber sorted i etrix probe se unique identi icial gene syr unique identi icial gene syr	176 NR3C1 wait L4821. 21. le testing (n=1732: seconds. Transmiti by strongest correla at identifier(s) of th at identifiers of gen fier for a sequence mbol ion about the varia ion coefficient ected p-value rate calculated by t <u>AccNum</u> NM_018695 NM_018183 NM_006218	() nuclear 7), please wait. ting 50 results: ation the gene used in es correlating t record [access] ble in the giver the Benjamini-F <u>GeneSym</u> ERBB2IP SBNO1 PIK3CA	the sear o the Lab ble at NC h row Hochberg Info 3 3 3	ch bel 1 gene BEI] Procedure Pearson r 0.924976 0.922925 0.922862	Raw p 4.449643e-45 1.693471e-44 1.762664e-44	Bonferroni 7.709896e-41 2.934277e-40 3.054167e-40	FDR (BH) 7.709896e-41 1.467138e-40 1.018056e-40
Fetchin Getting 1 of 1 c Calcula Done; 1 Colu E Colu E Calcula C Calcula C Calcula C Calcula C Calcula C Calcula C C C C C C C C C C C C C C C C C C C	I STITEST I g correlation I list of correlation I list of correlation whole program IIII of provention IIII of correction whole program IIIII of correction whole program IIIII of correction IIIII of correction IIIII of correction IIIII of correction IIIII of correction IIIIII of correction IIIIII of correction IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	data, please ations for id 1 es for id 148 ns for multip m took 0.793 end: mber sorted i etrix probe se unique identi icial gene syr unique identi icial gene syr	176 NR3C1 wait L4821. 21. le testing (n=1732: seconds. Transmiti by strongest correla ti dentifier(s) of th at identifiers of gen filer for a sequence nbol ion about the varia ion coefficient ected p-value rate calculated by t <u>AccNum</u> NM_018695 NM_018183 NM_006218 NM_022826	() nuclear 7), please wait. ting 50 results: ation the gene used in es correlating t record [access] ble in the giver the Benjamini-H ERBB2IP SBNO1 PIK3CA MARCH7	the sear o the Lab ble at NC h row Hochberg Info 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	ch bel 1 gene BEI] Procedure 0.924976 0.922925 0.922862 0.921074	Raw p 4.449643e-45 1.693471e-44 1.762664e-44 5.479506e-44	Bonferroni 7.709896e-41 2.934277e-40 3.054167e-40 9.494339e-40	FDR (BH) 7.709896e-41 1.467138e-40 1.018056e-40 2.373585e-40
Fetchin Getting 1 of 1 c Calcula Done; 1 Calcula Done; 1 Calcula L L L L L L L L L L L L L L L L L L L	I STITEST g correlation l list of correlation l list of correlation whole program Imm legge tank: Row nur abel: Affyrm abel: Affyrm abel: Affyrm ccNum: The I seneSym: Offi nfo: Link to m tearson r: Pea seneSym: Offi No (BH): Fals Label1 St14814 St14814 St14814 St14814 St14814 St14814	data, please ations for id : es for id 148 es for id 148 es for id 148 end: mber sorted l etrix probe se etrix	176 NR3C1 wait L4821. 21. le testing (n=1732: seconds. Transmiti by strongest correla at identifier(s) of th at identifiers of gen for a sequence nbol ion about the varia ion coefficient ected p-value rate calculated by t <u>AccNum</u> NM_018695 NM_018183 NM_002218 NM_022826 NM_015317	() nuclear 7), please wait. ting 50 results: ation the gene used in es correlating t record [accessi bile in the giver the Benjamini-H ERB82IP SBN01 PIK3CA MARCH7 PUM2	the sear o the Lab ble at NC h row Hochberg Info 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	ch bel 1 gene BEI] Procedure 0.924976 0.922925 0.922862 0.921074 0.920808	Raw p 4.449643e-45 1.693471e-44 1.762664e-44 5.479506e-44 6.470525e-44	Bonferroni 7.709896e-41 2.934277e-40 3.054167e-40 9.494339e-40 1.121148e-39	FDR (BH) 7.709896e-41 1.467138e-40 1.018056e-40 2.373585e-40 2.242296e-40
Fetchin Getting 1 of 1 c Calcula Done; 1 Calcula Done; 1 Calcula Calcu	g correlation list of correly done. list of correctio whole program mm legge tank: Row nur abel: Affynm abel: Affynm abel: Affynm abel: Affynm ccNum: The I senesym: Offi nfo: Link to m rearson r: Pea taw p: Unadju sonferroni: Bo DR (BH): Fals Label1 \$114814 \$114814 \$114814 \$114814 \$114814	data, please ations for id : es for id 148 es for id 148 es for id 148 end: mber sorted l atrix probe se atrix	176 NR3C1 wait L4821. 21. le testing (n=1732: seconds. Transmitti by strongest correlations ti dentifier(s) of this ti dentifiers of generic mbol solution about the varia ion coefficient ected p-value rate calculated by the NM_018695 NM_018183 NM_0015317 NM_003590	() nuclear 7), please wait. ting 50 results: ation the gene used in es correlating t record [accessi bile in the giver the Benjamini-H ERB82IP SBN01 PIK3CA MARCH7 PUM2 CUL3	the sear o the Lat ble at NO h row Hochberg Info 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	ch bel 1 gene BEI] Procedure 0.924976 0.922925 0.922862 0.921074 0.920808 0.921074	Raw p 4.449643e-45 1.693471e-44 1.762664e-44 5.479506e-44 6.470525e-44 9.136088e-43	Bonferroni 7.709896e-41 2.934277e-40 3.054167e-40 9.494339e-40 1.121148e-39 1.583010e-38	FDR (BH) 7.709896e-41 1.467138e-40 1.018056e-40 2.373585e-40 2.242296e-40 2.638350e-39
Fetchin Getting Collula Done; 1 Collula Collul	g correlation list of correly done. ting correctio whole program mm legge tank: Row nur abel: Affynm abel: Affynm abel: Affynm abel: Affynm ccNum: The I seneSym: Offi mfo: Link to m rearson r: Pea taw p: Unadju sonferroni: Bo DR (BH): Fals Label1 \$114814 \$114814 \$114814 \$114814 \$114814 \$114814 \$114814	data, please ations for id : es for id 148 es for id 148 es for id 148 end: mber sorted l atrix probe si etrix	176 NR3C1 wait L4821. 21. le testing (n=1732: seconds. Transmiti by strongest correla at identifier(s) of th ti dentifiers of gen mbol ion about the varia ion coefficient ected p-value rate calculated by th NM_018695 NM_018183 NM_00218 NM_022826 NM_015317 NM_003590 NM_012158	() nuclear 7), please wait. ting 50 results: ation the gene used in es correlating t record [access] ble in the giver the Benjamini-H ERB82IP SBN01 PIK3CA MARCH7 PUM2 CUL3 FBXL3	the sear o the Lat ble at NO h row Hochberg Info 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	ch bel 1 gene BEI 9 Procedure 0.924976 0.922925 0.922862 0.921074 0.920808 0.910449 0.916036	Raw p 4.449643e-45 1.693471e-44 1.762664e-44 5.479506e-44 6.470525e-44 9.136088e-43 1.165136e-42	Bonferroni 7.709896e-41 2.934277e-40 3.054167e-40 9.494339e-40 1.121148e-39 1.583010e-38 2.018830e-38	FDR (BH) 7.709896e-41 1.467138e-40 1.018056e-40 2.373585e-40 2.42296e-40 2.43350e-39 2.884043e-39
1492 Fetchin Getting 1 of 1 c Calcula Done; 1 • </td <td>g correlation list of correla- ting correction whole program mm legge tank: Row nur abel1: Affyme abel2: Affyme abel2: Affyme abel2: Affyme carson r: Pea aber y Unadju sonferroni: Bo DR (BH): Fals Label1 8114814 8114814 8114814 8114814 8114814 8114814 8114814 8114814 8114814 8114814</td> <td>data, please ations for id : es for id 148 es for id 148 es for id 148 es for id 148 end: mber sorted l etrix probe si etrix p</td> <td>176 NR3C1 wait L4821. 21. le testing (n=1732: seconds. Transmiti by strongest correla at identifier(s) of th at identifiers of gen mbol ion about the varial ion coefficient ected p-value rate calculated by th NM_018695 NM_018183 NM_00218 NM_022826 NM_015317 NM_003590 NM_012158 NM_022079</td> <td>() nuclear 7), please wait. ting 50 results: ation e gene used in es correlating t record [accessi ble in the giver the Benjamini-F ERBB2IP ERBB2IP SBN01 PIK3CA MARCH7 PUM2 CUL3 FBXL3 HERC4</td> <td>the sear o the Lat ble at NO h row Hochberg Info 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1</td> <td>ch bel 1 gene BIJ Procedure 0.924976 0.922925 0.922862 0.922862 0.921074 0.920808 0.921074 0.920808 0.916403 0.916036 0.914671</td> <td>Raw p 4.449643e-45 1.693471e-44 1.762664e-44 5.479506e-44 6.470525e-44 9.136088e-43 1.165136e-42 2.580860e-42</td> <td>Bonferroni 7.709896e-41 2.934277e-40 3.054167e-40 1.121148e-39 1.583010e-38 2.018830e-38 4.471856e-38</td> <td>FDR (BH) 7.709896e-41 1.467138e-40 1.018056e-40 2.373585e-40 2.242296e-40 2.638350e-39 2.884043e-39 5.589820e-39</td>	g correlation list of correla- ting correction whole program mm legge tank: Row nur abel1: Affyme abel2: Affyme abel2: Affyme abel2: Affyme carson r: Pea aber y Unadju sonferroni: Bo DR (BH): Fals Label1 8114814 8114814 8114814 8114814 8114814 8114814 8114814 8114814 8114814 8114814	data, please ations for id : es for id 148 es for id 148 es for id 148 es for id 148 end: mber sorted l etrix probe si etrix p	176 NR3C1 wait L4821. 21. le testing (n=1732: seconds. Transmiti by strongest correla at identifier(s) of th at identifiers of gen mbol ion about the varial ion coefficient ected p-value rate calculated by th NM_018695 NM_018183 NM_00218 NM_022826 NM_015317 NM_003590 NM_012158 NM_022079	() nuclear 7), please wait. ting 50 results: ation e gene used in es correlating t record [accessi ble in the giver the Benjamini-F ERBB2IP ERBB2IP SBN01 PIK3CA MARCH7 PUM2 CUL3 FBXL3 HERC4	the sear o the Lat ble at NO h row Hochberg Info 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	ch bel 1 gene BIJ Procedure 0.924976 0.922925 0.922862 0.922862 0.921074 0.920808 0.921074 0.920808 0.916403 0.916036 0.914671	Raw p 4.449643e-45 1.693471e-44 1.762664e-44 5.479506e-44 6.470525e-44 9.136088e-43 1.165136e-42 2.580860e-42	Bonferroni 7.709896e-41 2.934277e-40 3.054167e-40 1.121148e-39 1.583010e-38 2.018830e-38 4.471856e-38	FDR (BH) 7.709896e-41 1.467138e-40 1.018056e-40 2.373585e-40 2.242296e-40 2.638350e-39 2.884043e-39 5.589820e-39
1492 Fetchin Getting Colu Done; 1 of 1 c Calcula Done; * L L *	g correlation list of correlation list of correlation done. list of correlation whole program mn legge tank: Row nur abel1: Affyme abel2: Affyme abel2: Affyme abel2: Affyme correction: Bo DR (BH): Fals Label1 8114814 8114814 8114814 8114814 8114814 8114814 8114814 8114814 8114814 8114814 8114814 8114814 8114814 8114814 8114814 8114814	data, please ations for id : es for id 148 es for with the set es for correlat sted p-value unique identificial gene syr nore informati cial gene syr society socity society society socity society society	176 NR3C1 wait L4821. 21. le testing (n=1732: seconds. Transmiti by strongest correla ti dentifier(s) of th at identifiers of gen mbol ion about the varia ion coefficient ected p-value rate calculated by th NM_018695 NM_018183 NM_00218 NM_00218 NM_0012158 NM_012158 NM_022079 NM_144658	nuclear n	the sear o the Lat ble at NO h row Hochberg	ch bel 1 gene BI Procedure 0.924976 0.922925 0.922862 0.922862 0.922862 0.922074 0.922862 0.922862 0.921074 0.920808 0.916449 0.916036 0.914671 0.914475	Raw p 4.449643e-45 1.693471e-44 1.762664e-44 5.479506e-44 6.470525e-44 9.136083e-43 1.65136e-42 2.580860e-42 2.890279e-42	Bonferroni 7.709896e-41 2.934277e-40 3.054167e-40 9.494339e-40 1.121148e-39 1.533010e-38 2.018830e-38 4.471856e-38 5.007986e-38	FDR (BH) 7.709896e-41 1.467138e-40 1.018056e-40 2.373585e-40 2.42296e-40 2.43350e-39 2.884043e-39 5.589820e-39 5.564429e-39
1492 Fetchin Getting Colu Done; 1 of 1 c Calcula Done; P R L L L Calcula Done; R G J R S F R S G J <t< td=""><td>I STITEST g correlation list of correlation list of correlation list of correlation whole program Imm legge tank: Row nur abel1: Affyrme abel2: Affyrme abel2: Affyrme cellum in the relation cellum in the relation to the relation t</td><td>data, please ations for id : es for id 148 es for id 148 e</td><td>176 NR3C1 wait L4821. 21. le testing (n=1732: seconds. Transmitti by strongest correla ti dentifier(s) of th ti dentifiers of gen mbol ion about the varia ion coefficient ected p-value rate calculated by ti <u>AccNum</u> NM_018695 NM_018183 NM_006218 NM_0022826 NM_01317 NM_001317 NM_0012518 NM_022079 NM_144658 NM_001025105</td><td>() nuclear 7), please wait. ting 50 results: ation e gene used in es correlating t record [accessi ble in the giver the Benjamini-F ERBB2IP SBN01 PIK3CA MARCH7 PUM2 CUL3 FBXL3 HERC4 DOCK11 CSNK1A1</td><td>the sear o the Lat ble at NO h row Hochberg</td><td>ch bel 1 gene BI Procedure 0.924976 0.922925 0.922925 0.922925 0.922925 0.922925 0.922925 0.922925 0.922925 0.922925 0.921074 0.922925 0.921074 0.922925 0.921074 0.922925 0.921074 0.921074 0.91164</td><td>Raw p 4.449643e-45 1.693471e-44 1.762664e-44 6.470525e-44 9.136083e-43 1.165136e-42 2.580860e-42 2.890279e-42 1.876191e-41</td><td>Bonferroni 7.709896e-41 2.934277e-40 3.054167e-40 9.494339e-40 1.121148e-39 1.533010e-38 2.018830e-38 4.471858e-38 5.007986e-31 3.250877e-37</td><td>FDR (BH) 7.709896e-41 1.467138e-40 1.018056e-40 2.373585e-40 2.242296e-40 2.463350e-39 2.884043e-39 5.58820e-39 5.56429e-39 3.250877e-38</td></t<>	I STITEST g correlation list of correlation list of correlation list of correlation whole program Imm legge tank: Row nur abel1: Affyrme abel2: Affyrme abel2: Affyrme cellum in the relation cellum in the relation to the relation t	data, please ations for id : es for id 148 es for id 148 e	176 NR3C1 wait L4821. 21. le testing (n=1732: seconds. Transmitti by strongest correla ti dentifier(s) of th ti dentifiers of gen mbol ion about the varia ion coefficient ected p-value rate calculated by ti <u>AccNum</u> NM_018695 NM_018183 NM_006218 NM_0022826 NM_01317 NM_001317 NM_0012518 NM_022079 NM_144658 NM_001025105	() nuclear 7), please wait. ting 50 results: ation e gene used in es correlating t record [accessi ble in the giver the Benjamini-F ERBB2IP SBN01 PIK3CA MARCH7 PUM2 CUL3 FBXL3 HERC4 DOCK11 CSNK1A1	the sear o the Lat ble at NO h row Hochberg	ch bel 1 gene BI Procedure 0.924976 0.922925 0.922925 0.922925 0.922925 0.922925 0.922925 0.922925 0.922925 0.922925 0.921074 0.922925 0.921074 0.922925 0.921074 0.922925 0.921074 0.921074 0.91164	Raw p 4.449643e-45 1.693471e-44 1.762664e-44 6.470525e-44 9.136083e-43 1.165136e-42 2.580860e-42 2.890279e-42 1.876191e-41	Bonferroni 7.709896e-41 2.934277e-40 3.054167e-40 9.494339e-40 1.121148e-39 1.533010e-38 2.018830e-38 4.471858e-38 5.007986e-31 3.250877e-37	FDR (BH) 7.709896e-41 1.467138e-40 1.018056e-40 2.373585e-40 2.242296e-40 2.463350e-39 2.884043e-39 5.58820e-39 5.56429e-39 3.250877e-38
1492 Fetchin Getting Colu Done; *	I STITEST g correlation list of correlation list of correlation list of correlation whole program Imm legge tank: Row nur abel1: Affyrme abel2: Affyrme cell abel1: Affyrme abel2: Affyrme cell abel1: Affyrme abel2: Affyrme beneSym: Offi beneSym: O	data, please ations for id : es for id 148 ans for multip m took 0.793 and: mber sorted i etrix probe sis etrix probe sis etri	176 NR3C1 wait L4821. 21. le testing (n=1732: seconds. Transmitti by strongest correla et identifier(s) of th at identifier(s) of th ti dentifiers of gen mbol ion about the varial ion coefficient ected p-value rate calculated by ti <u>AccNum</u> NM_018695 NM_018183 NM_006218 NM_0022826 NM_013177 NM_00125105 NM_0125105 NM_0125105 NM_175854	() nuclear 7), please wait. ting 50 results: e gene used in es correlating t record [accessi ble in the giver the Benjamini-P ERBB2IP SBN01 PIK3CA MARCH7 PUM2 CUL3 FBXL3 HERC4 DOCK11 CSNK1A1 PAN3	the sear o the Lat ble at NO h row Hochberg	ch bel 1 gene BI Procedure 0.924976 0.922925 0.922925 0.922925 0.922925 0.922925 0.922925 0.922925 0.922925 0.922925 0.921074 0.920808 0.910479 0.914671 0.914671 0.914475 0.911164 0.911027	Raw p 4.449643e-45 1.693471e-44 1.762664e-44 6.470525e-44 9.136088e-43 1.165136e-42 2.580860e-42 2.890279e-42 1.876191e-41 2.023974e-41	Bonferroni 7.709896e-41 2.934277e-40 3.054167e-40 1.121148e-39 1.533010e-38 2.018830e-38 2.018830e-38 3.0507986e-31 3.506798e-38 3.250877e-37 3.506940e-37	FDR (BH) 7.709896e-41 1.467138e-40 1.018056e-40 2.373585e-40 2.242296e-40 2.463350e-39 2.884043e-39 5.58820e-39 5.56429e-39 3.250877e-38 3.188127e-38
1492 Fetchin Getting I of 1 c Calcula Done; 1 • </td <td>g correlation list of correlation list of correlation list of correlation whole program mn legge tank: Row nur abel1: Affyme abel2: Affyme abel2: Affyme collar the second abel1: Affyme abel2: Affyme abel2: Affyme beneSym: Off beneSym: Off beneSym be</td> <td>data, please ations for id : es for id 148 sns for multip m took 0.793 end: mber sorted i etrix probe sis etrix probe sis etri</td> <td>176 NR3C1 wait L4821. 21. le testing (n=1732: seconds. Transmitti by strongest correla et identifier(s) of th a sequence mbol no coefficient ected p-value rate calculated by ti <u>AccNum</u> NM_018695 NM_018183 NM_006218 NM_006218 NM_0022826 NM_015317 NM_0015317 NM_0015317 NM_00125105 NM_0125105 NM_0125105 NM_175854 NR_023350</td> <td>() nuclear 7), please wait. ting 50 results: e gene used in es correlating t record [accessi ble in the giver the Benjamini-P ERBB2IP SBN01 PIK3CA MARCH7 PUM2 CUL3 FBXL3 HERC4 DOCK11 CSNK1A1 PAN3 COPB2</td> <td>the sear o the Lat ble at NO hochberg Info 10 10 10 10 10 10 10 10 10 10 10 10 10</td> <td>ch bel 1 gene BI Procedure Pearson r 0.924976 0.922925 0.922925 0.922925 0.922925 0.922925 0.922925 0.922925 0.922925 0.921074 0.920808 0.921074 0.910630 0.91164 0.911027 0.9110830</td> <td>Raw p 4.449643e-45 1.693471e-44 1.762664e-44 5.479506e-44 6.470525e-44 9.136088e-43 1.165136e-42 2.580860e-42 2.890279e-42 1.876191e-41 2.023974e-41 2.256134e-41</td> <td>Bonferroni 7.709896e-41 2.934277e-40 3.054167e-40 1.121148e-39 1.121148e-39 2.018830e-38 2.018830e-38 3.05940e-37 3.506940e-37 3.909203e-37</td> <td>FDR (BH) 7.709896e-41 1.467138e-40 1.018056e-40 2.373585e-40 2.242296e-40 2.373585e-40 2.373585e-40 2.538320e-39 5.568429e-39 3.250877e-38 3.188127e-38 3.257669e-38</td>	g correlation list of correlation list of correlation list of correlation whole program mn legge tank: Row nur abel1: Affyme abel2: Affyme abel2: Affyme collar the second abel1: Affyme abel2: Affyme abel2: Affyme beneSym: Off beneSym: Off beneSym be	data, please ations for id : es for id 148 sns for multip m took 0.793 end: mber sorted i etrix probe sis etrix probe sis etri	176 NR3C1 wait L4821. 21. le testing (n=1732: seconds. Transmitti by strongest correla et identifier(s) of th a sequence mbol no coefficient ected p-value rate calculated by ti <u>AccNum</u> NM_018695 NM_018183 NM_006218 NM_006218 NM_0022826 NM_015317 NM_0015317 NM_0015317 NM_00125105 NM_0125105 NM_0125105 NM_175854 NR_023350	() nuclear 7), please wait. ting 50 results: e gene used in es correlating t record [accessi ble in the giver the Benjamini-P ERBB2IP SBN01 PIK3CA MARCH7 PUM2 CUL3 FBXL3 HERC4 DOCK11 CSNK1A1 PAN3 COPB2	the sear o the Lat ble at NO hochberg Info 10 10 10 10 10 10 10 10 10 10 10 10 10	ch bel 1 gene BI Procedure Pearson r 0.924976 0.922925 0.922925 0.922925 0.922925 0.922925 0.922925 0.922925 0.922925 0.921074 0.920808 0.921074 0.910630 0.91164 0.911027 0.9110830	Raw p 4.449643e-45 1.693471e-44 1.762664e-44 5.479506e-44 6.470525e-44 9.136088e-43 1.165136e-42 2.580860e-42 2.890279e-42 1.876191e-41 2.023974e-41 2.256134e-41	Bonferroni 7.709896e-41 2.934277e-40 3.054167e-40 1.121148e-39 1.121148e-39 2.018830e-38 2.018830e-38 3.05940e-37 3.506940e-37 3.909203e-37	FDR (BH) 7.709896e-41 1.467138e-40 1.018056e-40 2.373585e-40 2.242296e-40 2.373585e-40 2.373585e-40 2.538320e-39 5.568429e-39 3.250877e-38 3.188127e-38 3.257669e-38
Instruct Fetchin Getting Getting I of 1 c Calcula Done; 1 Fetchin R L L L E Fetchin R Getting P R Getting P R Getting R S Getting Getting Getting R Getting Getting Getting Getting Getting Getting Getting Geting Geting	I STITEST g correlation list of correlation list of correlation whole program Imm legge tank: Row nur abel: Affyma abel2: Affyma	data, please ations for id : es for id 148 ses for id 148 es for id 148	176 NR3C1 wait L4821. 21. le testing (n=1732: seconds. Transmitt by strongest correla et identifier(s) of th at identifiers of gen bol iden about the varia ion coefficient ected p-value rate calculated by th NM_018183 NM_006218 NM_015317 NM_0022826 NM_015317 NM_001517 NM_0012518 NM_00125105 NM_175854 NR_023350 NM_004521 	nuclear	the sear o the Lat ble at No hochberg Info 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	ch bel 1 gene BI Procedure Pearson r 0.924976 0.922925 0.922925 0.922862 0.922062 0.921074 0.920808 0.91074 0.910636 0.914671 0.914671 0.91164 0.911027 0.9110830 0.910830 0.909288	Raw p 4.449643e-45 1.693471e-44 1.762664e-44 6.470525e-44 6.470525e-44 1.165136e-42 2.580860e-42 2.580860e-42 2.890279e-42 1.876191e-41 2.255134e-41 5.243304e-41	Bonferroni 7.709896e-41 2.934277e-40 3.054167e-40 3.054167e-40 9.494339e-40 1.121148e-39 1.583010e-38 2.018830e-38 2.018830e-38 2.018830e-38 3.506940e-37 3.506940e-37 3.506940e-37 3.909203e-37	FDR (BH) 7.709896e-41 1.467138e-40 1.467138e-40 2.373585e-40 2.242296e-40 2.373585e-40 2.584043e-39 5.589820e-39 5.564429e-39 3.188127e-38 3.188127e-38 3.257669e-38 6.988517e-38

general functions were attributed to the genes most strongly associated with GR. Furthermore, in the intercorrelation network generated by IPA, all the tested transcriptional regulators (CREB1, GATA1, and GR) had a very central position in the respective top ranked networks (Figs. S1, S2 and S3), supporting that the detected associations were real.

Discussion

We hypothesized that highly correlated blood cells transcripts could be functionally associated in our dataset and that these associations could be easily assessed by a user-friendly web application. We explored significant macromolecular associations involving CREB1, GATA1 and GR transcription factors. The in silico associations were supported using ENCODE ChIP data from both tissues and cells unrelated to blood, indicating a common functionality irrespective of cell or tissue type. The finding that fewer significant GR correlations (71/200, Table S3) were identified in ENCODE compared to CREB1 (182/200, Table S1) and GATA1 (146/200, Table S2) may be related to GR being able to bind other transcription factors, e.g., those binding to The Activator Protein-1 (AP-1) sites (Fos, Jun and others) without binding directly to DNA [25]. For example, only 62% of dexamethasone induced GR binding sites contained the GR response element when dexamethasone induced transcription was studied in A549 cells [26]. Since thousands of human genes harbour DNA binding elements for the tested transcription factors, we expected to find such elements also in several randomly tested genes used as control, but significantly less. The very high overlap in transcript association between the Framingham and Oslo cohorts, confirmed the validity of the results obtained using in silico analyses. Correlation estimates from the Oslo cohort were generally somewhat higher than in the Framingham dataset, and the difference may be related to a more heterogenous Framingham cohort with respect to age, ethnicity and health status. As expected, experimentally verified mRNA/miRNA associations were not reproduced equally well as mRNA/ mRNA associations in our data. This is probably because cell/tissue specific sets of miRNAs are needed to target and degrade mRNAs [2]. Sets of miRNAs targeting specific mRNAs identified in other cohorts and/or tissues may not be present in peripheral blood. Thus, we consider replication of 25% (Table S4) and 43% (Table S5) of verified miRNA/mRNA interactions as satisfactory. The results underscore the usefulness of the in silico approach and web application for detection of miRNA/ mRNA associations in peripheral blood, but appear also to have relevance for other tissues. We assume that associations identified are relevant for both sexes, but this needs to be verified.

Conclusion

The results indicate that in silico analyses using a large correlation matrix containing the blood transcriptome associations in combination with a user-friendly web application, may identify functionally associated macromolecules in blood with relevance also for tissues.

Supplementary Information

The online version contains supplementary material available at https://doi. org/10.1186/s12863-022-01077-3.

Additional file 1.

Additional file 2.

Acknowledgements

Not applicable.

Authors' contributions

K.M.G., S.R., and C.L. jointly supervised this work. S.R. designed and managed individual studies. D.S. designed the application. K.M.G., S.R., Y-H. S., and D.K. collected data. K.M.G., S.R., and C.L. reviewed the analysis plan. S.R., O.K.O., and A.C.H. analyzed the data. S.R. supervised the overall study design. K.M.G., T.P.U., C.L., and S.R. wrote the manuscript which was reviewed by all authors. The author (s) read and approved the final manuscript.

Funding

This work was supported by the South East Norway Health Authority and Oslo University Hospital, Ullevaal (52009/8029); The 6th EU Framework Program (LSHM-CT-2003-502941); Legat til Forskning, Lovisenberg Diaconal Hospital; The National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health (R01 AR041398 and R01 AR072199) Kiel has received grants to his institution from Amgen, Inc., and Radius Health.

Availability of data and materials

The normalized PCR and Affymetrix signal values used for generation of data accessible in the web application are available in Table S7. The application is available at: http://app.uio.no/med/klinmed/correlation-browser/blood/index.php The primary data used in the web application have also been deposited in NCBI's Gene Expression Omnibus [27], accessible through GEO Series accession number GSE198941 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi? acc=GSE198941).

Declarations

Ethics approval and consent to participate

Prior to the study, approval was obtained by the Norwegian Regional Ethics Committee (REK 2010/2539) for the Oslo cohort and by Boston University Medical Center Institutional Review Board for the Framingham cohort. The study was conducted according to the Declaration of Helsinki (2000). Sampling and procedures followed the National laws after obtaining verbal and written, informed consent. The verbal informed consent obtained from the participants were approved by the Norwegian Regional Ethics Committee (REK 2010/2539) and Boston University Medical Center Institutional Review Board. Validation data was made available from dbGaP through approved request number 1302685–1.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Unger-Vetlesen Institute, Lovisenberg Diaconal Hospital, Lovisenberggata 17, 0456 Oslo, Norway. ²Department of Medical Biochemistry, Oslo University Hospital, Oslo, Norway. ³Maine Medical Center Research Institute, Scarborough, ME, USA. ⁴Department of Medicine, The Hinda and Arthur Marcus Institute for Aging Research, Beth Israel Deaconess Medical Center and Harvard Medical School and Broad Institute of MIT & Harvard, Boston, MA, USA. ⁵Department of Plastic and Reconstructive Surgery, Oslo University Hospital, Oslo, Norway. ⁶Department of Ophthalmology, Stavanger University Hospital, Arendal, Arendal, Norway.

Received: 2 March 2022 Accepted: 13 July 2022 Published online: 26 July 2022

References

- 1. Levine M, Tjian R. Transcription regulation and animal diversity. Nature. 2003;424(6945):147–51.
- Saliminejad K, Khorram Khorshid HR, Soleymani Fard S, Ghaffari SH. An overview of microRNAs: biology, functions, therapeutics, and analysis methods. J Cell Physiol. 2019;234(5):5451–65.
- Segal E, Shapira M, Regev A, Pe'er D, Botstein D, Koller D, et al. Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. Nat.Genet. 2003;34(2):166–76.
- Zeng T, Li J. Maximization of negative correlations in time-course gene expression data for enhancing understanding of molecular pathways. Nucleic Acids Res. 2010;38(1):e1.
- Wang H, Wang Q, Pape UJ, Shen B, Huang J, Wu B, et al. Systematic investigation of global coordination among mRNA and protein in cellular society. BMC Genomics. 2010;11:364.
- Wen AY, Sakamoto KM, Miller LS. The role of the transcription factor CREB in immune function. J Immunol. 2010;185(11):6413–9.
- Wang P, Tian H, Zhang J, Qian J, Li L, Shi L, et al. Spaceflight/microgravity inhibits the proliferation of hematopoietic stem cells by decreasing kit-Ras/cAMP-CREB pathway networks as evidenced by RNA-Seq assays. FASEB J. 2019;33(5):5903–13.
- Liu L, Karmakar S, Dhar R, Mahajan M, Choudhury A, Weissman S, et al. Regulation of Ggamma-globin gene by ATF2 and its associated proteins through the cAMP-response element. PLoS One. 2013;8(11):e78253.
- Sandoval S, Kraus C, Cho EC, Cho M, Bies J, Manara E, et al. Sox4 cooperates with CREB in myeloid transformation. Blood. 2012;120(1):155–65.
- Altarejos JY, Montminy M. CREB and the CRTC co-activators: sensors for hormonal and metabolic signals. NatRevMolCell Biol. 2011;12(3):141–51.
- Bartolotti N, Lazarov O. CREB signals as PBMC-based biomarkers of cognitive dysfunction: a novel perspective of the brain-immune axis. Brain Behav Immun. 2019;78:9–20.
- 12. Nayak RR, Kearns M, Spielman RS, Cheung VG. Coexpression network based on natural variation in human gene expression reveals gene interactions and functions. Genome Res. 2009;19(11):1953–62.
- Beck IM, De BK, Haegeman G. Glucocorticoid receptor mutants: manmade tools for functional research. Metab: Trends Endocrinol; 2011.
- Hofbauer LC, Rauner M. Minireview: live and let die: molecular effects of glucocorticoids on bone cells. Mol.Endocrinol. 2009;23(10):1525–31.
- Tait AS, Butts CL, Sternberg EM. The role of glucocorticoids and progestins in inflammatory, autoimmune, and infectious disease. J.Leukoc.Biol. 2008;84(4):924–31.
- Reppe S, Refvem H, Gautvik VT, Olstad OK, Hovring PI, Reinholt FP, et al. Eight genes are highly associated with BMD variation in postmenopausal Caucasian women. Bone. 2010;46(3):604–12.
- Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP. An investigation of coronary heart disease in families. The Framingham offspring study, Am J Epidemiol. 1979;110(3):281–90.
- Joehanes R, Johnson AD, Barb JJ, Raghavachari N, Liu P, Woodhouse KA, et al. Gene expression analysis of whole blood, peripheral blood mononuclear cells, and lymphoblastoid cell lines from the Framingham heart study. Physiol Genomics. 2012;44(1):59–75.
- Joehanes R, Ying S, Huan T, Johnson AD, Raghavachari N, Wang R, et al. Gene expression signatures of coronary heart disease. Arterioscler Thromb Vasc Biol. 2013;33(6):1418–26.

- Reppe S, Sachse D, Olstad OK, Gautvik VT, Sanderson P, Datta HK, et al. Identification of transcriptional macromolecular associations in human bone using browser based in silico analysis in a giant correlation matrix. Bone. 2012;53(1):69–78.
- Benjamini Y, Hochberg Y. Controlling the false discovery rate a practical and powerful approach to multiple testing. J Royal Stat Soc Series B-Methodolo. 1995;57(1):289–300.
- 22. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. Ann Stat. 2001;29(4):1165–88.
- Rouillard AD, Gundersen GW, Fernandez NF, Wang Z, Monteiro CD, McDermott MG, et al. The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. Database (Oxford). 2016;2016.
- Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, Vlachos IS, Tastsoglou S, Kanellos I, et al. DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA-gene interactions. Nucleic Acids Res. 2018;46(D1):D239–45.
- Herrlich P. Cross-talk between glucocorticoid receptor and AP-1. Oncogene. 2001;20(19):2465–75.
- Reddy TE, Pauli F, Sprouse RO, Neff NF, Newberry KM, Garabedian MJ, et al. Genomic determination of the glucocorticoid response reveals unexpected mechanisms of gene regulation. Genome Res. 2009;19(12):2163–71.
- Edgar R, Domrachev M, Lash AE. Gene expression omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002;30(1):207–10.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

