Systems-level approaches reveal conservation of trans-regulated genes in the rat and genetic determinants of blood pressure in humans

Sarah R. Langley1, Leonardo Bottolo2, Jaroslav Kunes3, Josef Zicha3, Vaclav Zidek3, Norbert Hubner4, Stuart A. Cook1,5, Michal Pravenec3,6, Timothy J. Aitman1, and Enrico Petretto1*

1MRC Clinical Sciences Centre, Faculty of Medicine, Imperial College London, Du Cane Road, London W12 0NN, UK; 2Department of Mathematics, Imperial College London, London SW7 2AZ, UK; 3Institute of Physiology, Academy of Sciences of the Czech Republic, Prague 142 20, Czech Republic; 4Max-Delbruck-Center for Molecular Medicine, Berlin, Germany; 5Heart and Lung Institute, Imperial College, Dovehouse Street, London SW3 6LY, UK; and 6Institute of Biology and Medical Genetics, 1st Medical Faculty, Charles University, Prague 12800, Czech Republic

Received 4 July 2012; revised 8 October 2012; accepted 26 October 2012; online publish-ahead-of-print 31 October 2012

Aims

Human genome-wide association studies (GWAS) of hypertension identified only few susceptibility loci with large effect that were replicated across populations. The vast majority of genes detected by GWAS has small effect and the regulatory mechanisms through which these genetic variants cause disease remain mostly unclear. Here, we used comparative genomics between human and an established rat model of hypertension to explore the transcriptional mechanisms mediating the effect of genes identified in 15 hypertension GWAS.

Methods and results

Time series analysis of radiotelemetric blood pressure (BP) was performed to assess 11 parameters of BP variation in recombinant inbred strains derived from the spontaneously hypertensive rat. BP data were integrated with ≏27 000 expression quantitative trait loci (eQTLs) mapped across seven tissues, detecting 8000 significant associations between eQTL genes and BP variation in the rat. We then compiled a large catalogue of human genes from GWAS of hypertension and identified a subset of 2292 rat–human orthologous genes. Expression levels for 795 (34%) of these genes correlated with BP variation across rat tissues: 51 genes were cis-regulated, whereas 459 were trans-regulated and enriched for ‘calcium signalling pathway’ ($P = 9.6 \times 10^{-6}$) and ‘ion channel’ genes ($P = 3.5 \times 10^{-7}$), which are important determinants of hypertension. We identified 158 clusters of trans-eQTLs, annotated the underlying ‘master regulator’ genes and found significant over-representation in the human hypertension gene set (enrichment $P = 5 \times 10^{-4}$).

Conclusion

We showed extensive conservation of trans-regulated genes and their master regulators between rat and human hypertension. These findings reveal that small-effect genes associated with hypertension by human GWAS are likely to exert their action through coordinate regulation of pathogenic pathways.

Keywords

Integrative genomics • Expression QTLs • Time series analysis • Trans-acting regulation • Genome-wide association studies

1. Introduction

Hypertension, characterized by elevated blood pressure (BP), is one of the strongest risk factors for common cardiovascular diseases worldwide. Despite the high heritability of BP variation, human genome-wide association studies (GWAS) have been able to identify only few susceptibility loci with a large effect for hypertension that were replicated across populations and a large proportion of the underlying genetic determinants remains to be investigated. Moreover, the functional context and regulatory mechanisms through which the susceptibility variants identified by GWAS cause hypertension or elevated BP remain mostly unclear.
Common genetic variation underlying disease susceptibility might exert its function by altering gene expression in a specific tissue. Quantitative variation in gene expression that is due to underlying sequence variation can be mapped to the genome in segregating populations as expression quantitative trait loci (eQTLs). Genetic variants can affect the expression of genes that reside locally (i.e. cis-eQTLs) or remotely (i.e. trans-eQTLs), reflecting different regulatory mechanisms of transcription. In particular, trans-eQTLs are informative of downstream-regulated genes, which can be associated with disease susceptibility and hence might reveal pathophysiological pathways.

GWAS focus on the identification of significant SNP(s)-disease associations and pinpoint local (cis-acting) genetic effects. Similarly, typical GWAS have limited power to detect secondary (trans-acting) regulatory effects of the associated genes on disease or complex phenotypes. Gene expression can be considered as an intermediate ‘endo-phenotype’ between genetic variation and whole body disease phenotypes, and can be used to investigate the regulatory mechanisms mediating genetic susceptibility to disease. For example, genome-wide expression levels can be directly correlated with variation in pathophysiological phenotypes by quantitative trait transcript (QTT) analysis to pinpoint eQTL genes and transcripts associated with disease. This approach was successfully exploited in the rat model to identify several susceptibility genes for cardiovascular traits including BP variation and hypertension.

In this study, we investigate the regulatory mechanisms mediating the effect of the genetic associations with hypertension and BP variation identified in human GWAS, including both significant and replicated GWAS signals. To this aim we annotated a large catalogue of human hypertension genes we combined GWAS results from 15 large studies and adopted a relaxed threshold of significance (GWAS P-value) to ascertain associated genes (n = 4145), as previously described in Johnson and O’Donnell (see Section 2). Different gene sets in the rat where then tested for over-representation against the human-rat orthologues genes (n = 2292) associated with hypertension by GWAS.

2. Methods

2.1 Rat data sets

The BXH/HXB RI rat lines were derived by intercrossing BN and SHR strains as previously described. Genome-wide gene expression profiles are available in seven tissues using the Affymetrix 230 2.0 Array chip (aorta, liver, left ventricle, and skeletal muscle) and Affymetrix 230A Array chip (kidney, adrenal gland, and fat). Tissues were collected from four unfasted males (6 weeks old) of each RI strain between 9 a.m. and 10 a.m., and were frozen in liquid nitrogen and stored at –80°C. The gene expression levels were normalized using RMA and outliers within each RI strain removed using the Napliov test, as previously described. Time series blood pressure measurements were collected from 28 BXH/HXB RI lines. We implanted indwelling aortic radiotelemetry transducers (Data Sciences International) and measured arterial pressure in conscious, unrestrained rats. The Dataquest IV radiotelemetry system (Data Sciences International, St Paul, MN, USA) was used for the measurement of systolic, diastolic, mean arterial pressure, and HR. Transmitters (TA11PA) were implanted to 12-week-old males after the calibration as recommended by the supplier. Briefly, the rats were anaesthetized with ketamine (100 mg/kg) and xylazine (5 mg/kg), a midline abdominal incision was made and abdominal aorta was exposed. The adequacy of anaesthesia was assessed by testing the loss of the pedal reflex. The catheter of the transmitter was inserted rostrally into the aorta through a small hole close to the bifurcation and covered by the piece of nitrocellulose sheet to protect the bleeding. The body of the transmitter was sutured to the inside of abdominal muscle wall, whereas muscle and skin incisions were sutured close. Rats were housed individually in standard cages, which were placed over the receivers that were connected to the personal computer for data acquisition. The animals were left to recover for 10 days after the surgery. After the end of experiment, the rats killed by overdose of the above mentioned anaesthetics. Five hundred measurements per second were collected over a period of 30 min in each animal (recorded between 9 a.m. and 12 p.m.), and there are between 2 and 11 rats per strain. All of the animal studies were performed in agreement with the Animal Protection...
Law of the Czech Republic (311/1997) and approved by the Ethics Committee of the Czech Academy of Sciences Institute of Physiology. This investigation conforms to the Directive 2010/63/EU of the European Parliament.

2.2 Time series analysis

2.2.1 Conventional phenotype measurements
Diastolic blood pressure (DBP) was obtained by averaging the minimum peak values within an individual time series and systolic blood pressure (SBP) was similarly obtained by averaging the maximum peak values within an individual time series. The mean blood pressure (MBP) was obtained by averaging all blood pressure measurements. The heart rate was obtained by dividing the number of maximum peaks in an individual time series by the length of the time series, in seconds, to obtain a measurement in beats per second. Peak-to-peak distance (PDIST) was calculated by averaging the distance between maximum peaks. Pulse pressure (PP) was calculated by averaging the difference between adjacent maximum and minimum peaks.

2.2.2 Spectral phenotype measurements
To better characterize the cyclical and periodic nature of the blood pressure time series, we performed fast Fourier transform (FFT) analysis. Since the time series must be (weakly) stationary to obtain an accurate estimate of the spectral power, the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test was used to assess the existence of a trend and the level stationarity of each time series. In all time series analysed, the null hypothesis was rejected and a differencing method was used to correct for the non-stationarity. Each lagged time series was thinned by a factor of 10, so that the number of measurements was reduced from 500 to 50 to remove possible autocorrelation. All time series were then transformed from the time domain into the frequency domain via FFT. We defined SPEC as the height of the highest peak in the spectral power density of the original sequence, FREQ to be the frequency at which the peak occurs, and POWER is the average power calculated by integrating over the 95% confidence interval surrounding the peak. (For details, see Supplementary material online, Methods).

2.2.3 Wavelet phenotype measurements
Wavelet analysis allows to decompose a time series into time/frequency space simultaneously. First, the original time series with 500 measurements per second was filtered using a 1-D digital filter to reduce the trend in the time series. Secondly, a Morlet wavelet was applied and the local maximum of the wavelet coefficients was calculated within the HR range (4–8 Hz). Within this range, we defined WAVE as the frequency that corresponds to the local maximum peak and WFREQ as the frequency that corresponds to the HR range (4–8 Hz). Within this range, we defined WAVE as the frequency that corresponds to the local maximum peak and WFREQ as the frequency that corresponds to the HR range. Finally, for each of the conventional, spectral and wavelet phenotype, a Grubbs test was applied to identify and remove outliers within each strain; strains means were then used in the subsequent analysis (for details on the phenotypes, see Supplementary material online, Table S1). Conventional, spectral and wavelet phenotypes were clustered using partial correlation analysis by means of graphical Gaussian models, as previously described.

2.3 Expression QTL mapping
We employed a sparse Bayesian regression model to map eQTLs using genome-wide gene expression levels measured in the adrenal gland, aorta, fat, kidney, left ventricle, liver, and skeletal muscle tissues. A genome-wide SNP panel consisting of 1304 non-redundant SNPs was used and missing genotypes were imputed using FastPhase. Microarray probesets that map to multiple locations in the genome and non-expressed transcripts were removed prior to the analysis. This defined a set of 11 671 genes in the aorta, liver, left ventricle, skeletal muscle and 8444 genes in the adrenal, fat, kidney, respectively. Significance of the strength of transcript–SNP association (i.e. posterior probability of association (PPA)) was assessed using the method proposed in Chen et al. In detail, SNPs are ranked in a decreasing order according to their local false discovery rate (FDR) (i.e. 1-PPA) and the top-ranked SNPs with an average FDR < 10% are declared significant. For each microarray probeset whose expression mapped to a SNP (hence forming an eQTL), we defined a cis-eQTL when the location of the probeset is within 10 Mb of the physical location of the SNP. Trans-eQTLs were identified when the probeset was located > 10 Mb away from the SNP or on another chromosome. Trans-eQTL clusters were defined to be groups of 10 or more genes, which map in trans to the same location of the genome (i.e. the same SNP). Genes located within a 2 kb region around the SNP defining the trans-eQTL cluster were called putative ‘master regulators’ and represent candidate regulators of the trans-eQTLs. When multiple genes were found within the 2 kb region the gene nearest to the SNP defining the trans-cluster was prioritized.

2.4 Quantitative trait transcript analysis
To determine the association between blood pressure phenotypes and gene expression, we correlated each phenotype with genome-wide expression levels across seven tissues using pairwise Pearson correlation, as previously described. Using the Westfall-Young procedure, we corrected P-values of each pairwise correlation using 1000 permutations. Associations were deemed significant with an empirical P-value < 0.05.

2.5 Functional annotation analysis
The Kyoto Encyclopedia of Genes and Genomes pathway and gene ontology enrichment analysis was performed using DAVID. The significance of the enrichment was assessed by the Fisher’s exact test and correction for multiple testing was performed using Benjamini FDR.

2.6 Human genome-wide association studies and comparative genomics analysis
We used and expand the database of human GWAS results published by Johnson and O’Donnell to obtain a list of rat genes with human orthologues associated with hypertension. Specifically, the database contained all SNPs associated with many human diseases/traits from 118 GWAS published until 2008, and included SNPs that were significant based upon four conditions:

1. The SNP had an identifiable ID or verifiable genomic position.
2. A statistical P-value for association was reported.
3. The P-value was ≤ 0.001 (allowing for rounding) if the association was from a raw, unadjusted genome-wide association scan.
4. The P-value was ≤ 0.05 if the association was derived from replication, fine mapping, or re-sequence efforts, or if it was identified as belonging to a locus or region that was specifically identified as an a priori candidate by the authors.

We filtered the list of GWAS data to look only at hypertension related phenotypes (hypertension case–control, blood pressure, SBP, and DBP) resulting in five published studies. We expanded upon this list using the same criteria from GWAS published since 2008, which included an additional 10 research articles. The genes closest to, or encompassing, the human SNPs and the rat orthologues of those human genes were identified using Ensembl’s Biomart 67. We looked for an over-representation in the human GWAS genes for the following rat gene sets: (i) cis and trans-eQTLs mapped in seven tissues; (ii) non-eQTL genes; (iii) trans-eQTL clusters and genes underlying trans-eQTL clusters; (iv) the subset of each of the above (i–iii) that significantly correlated with any blood pressure trait by QTT analysis (FDR < 5%). Over-representation of rat genes in the set of human GWAS data was assessed using the hypergeometric test, and...
empirical P-values of significance were assessed using 10 000 permutations, with the size of the sample drawn equal to that of the set analysed.34

3. Results

3.1 Spectral and wavelet analysis of blood pressure variation in the rat

Radiotelemetric BP data were collected in the BXH/HXB RI strains panel14 at a frequency of 500 Hz over the course of 30 min.12 In addition to conventional approaches to measure standard BP indices (SBP, DBP, MBP, PP, PDIST, and HR), we used spectral and wavelet analysis to assess different physiological parameters of BP and accurately quantify variability of BP and HR signals. Spectral analysis transforms the BP time series from the time domain to the frequency domain, allowing the decomposition of the signal into periodic components that can be informative of variation of cardiovascular function and circadian BP rhythm.35 Similarly, wavelet analysis transforms the time series from the time domain to the frequency domain, retaining information about the periodic and variable components of BP and HR.36 We derived 11 phenotypes and showed that the new BP parameters (FREQ, POWER, SPEC, WAVE, and WFREQ) have heritability estimates that are comparable with the standard phenotypes (heritability range: 13–32%), suggesting a genetic component underlying the regulation of these traits (Table 1 and Supplementary material online, Table S1).

3.2 Extensive trans-acting regulation of gene expression across seven rat tissues

To identify primary and secondary regulators of BP variation in the rat, we integrated all BP parameters with genome-wide gene expression data across seven tissues (adrenal, aorta, fat, kidney, liver, left ventricle, and skeletal muscle) using sparse Bayesian regression methods,19 and identified 26 990 eQTLs (FDR <10%) across the seven tissues, both cis- and trans-acting (Figure 2 and Supplementary material online, Table S2). Overall, 4.0–7.0% of transcripts were found to be cis-regulated and 9.8–15.1% were found to be trans-regulated across tissues. In keeping with previous data,19,33 the stronger associations were observed for the cis-eQTLs (average marginal PPA = 0.94) when compared with the trans-eQTLs (average marginal PPA = 0.88), and on average we found 2.3-fold more trans- than cis-eQTLs (Figure 2). We looked at previously mapped blood pressure QTLs in SHR and SHR/BN strains (Ren et al., 2010) and found that 3713 eQTLs (~15% of all eQTLs identified) were located within these QTL regions (Supplementary material online, Table S3). In particular, 2439 (66%) of these eQTLs were trans-regulated, suggesting a role for trans-eQTLs in the regulation of known BP loci in the rat.

Multiple trans-regulated genes can map to the same locus forming clusters of eQTLs (or hot spots),38 which can provide direct insights into the co-ordinated regulatory mechanisms underlying disease.9,10 At FDR <10% we identified 158 trans-eQTL clusters across the seven tissues (Figure 3 and Supplementary material online, Table S4), where we defined a trans-eQTL cluster to be a group of 10 or more genes whose mRNA expression map in trans to the same genetic locus. Fourteen trans-eQTL clusters were detected in multiple tissues (Figure 3), suggesting common regulatory mechanisms that are conserved between multiple organs and systems. Since 14 clusters were detected in multiple tissues and pinpointed to the same QTL regions (Figure 3), only 138 trans-eQTL cluster loci mapped uniquely to the genome. One of these ‘replicated’ trans-eQTL clusters, detected in the kidney (kid15_106), liver (liver15_106), and left ventricle (lv15_106) tissues (Figure 3), was identified in previous studies and harboured a ‘master regulator’ gene at the locus, which was conserved between rats and humans.9 Here, we have systematically annotated all genes located at (or nearby) the SNP that define the trans-eQTL cluster loci identified across tissues (Figure 3). For each unique trans-eQTL cluster, we conservatively identified the closest gene to SNP defining the trans-eQTL cluster using a 2 kb region centred on the SNP marker. These genes were called putative ‘master regulators’ for the trans-eQTL clusters. Fifty-six of the 138 unique trans-eQTL clusters loci had an annotated gene within the 2 kb region and 12 of these genes were cis-regulated (Supplementary material online, Table S5).

<table>
<thead>
<tr>
<th>Trait name</th>
<th>Abbreviation</th>
<th>Type</th>
<th>Grouping</th>
<th>Median</th>
<th>Range</th>
<th>Heritability (h²) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diastolic blood pressure</td>
<td>DBP</td>
<td>Conventional</td>
<td>BP</td>
<td>80.55</td>
<td>59.03–94.63</td>
<td>21</td>
</tr>
<tr>
<td>Heart rate</td>
<td>HR</td>
<td>Conventional</td>
<td>Rate</td>
<td>298.5</td>
<td>276.3–352.4</td>
<td>23</td>
</tr>
<tr>
<td>Mean blood pressure</td>
<td>MBP</td>
<td>Conventional</td>
<td>BP</td>
<td>91.44</td>
<td>68.44–108.61</td>
<td>24</td>
</tr>
<tr>
<td>Peak-to-peak distance</td>
<td>pDist</td>
<td>Conventional</td>
<td>Rate</td>
<td>0.20</td>
<td>0.16–0.22</td>
<td>21</td>
</tr>
<tr>
<td>Pulse pressure</td>
<td>PP</td>
<td>Conventional</td>
<td>Pressure</td>
<td>32.82</td>
<td>22.16–40.49</td>
<td>21</td>
</tr>
<tr>
<td>Systolic blood pressure</td>
<td>SBP</td>
<td>Conventional</td>
<td>BP</td>
<td>112.92</td>
<td>87.26–142.93</td>
<td>31</td>
</tr>
<tr>
<td>Maximum frequency</td>
<td>Freq</td>
<td>Spectral</td>
<td>Rate</td>
<td>4.79</td>
<td>4.41–5.50</td>
<td>21</td>
</tr>
<tr>
<td>Total power</td>
<td>Power</td>
<td>Spectral</td>
<td>Pressure</td>
<td>67.03</td>
<td>28.66–133.18</td>
<td>25</td>
</tr>
<tr>
<td>Maximum spectra</td>
<td>Spec</td>
<td>Spectral</td>
<td>Pressure</td>
<td>0.17</td>
<td>0.10–0.44</td>
<td>18</td>
</tr>
<tr>
<td>Wave amplitude</td>
<td>Wave</td>
<td>Wavelet</td>
<td>Wave</td>
<td>0.25</td>
<td>0.08–0.49</td>
<td>13</td>
</tr>
<tr>
<td>Wave frequency</td>
<td>Wfreq</td>
<td>Wavelet</td>
<td>Rate</td>
<td>4.88</td>
<td>4.34–5.84</td>
<td>19</td>
</tr>
</tbody>
</table>

*Grouping was determined by clustering the blood pressure traits using partial correlations (FDR <5%), which were estimated by graphical Gaussian models (see Section 2).

For each blood pressure trait the narrow sense heritability (h²) in the RI strains was calculated, as previously described.27
3.3 Identification of rat genes associated with blood pressure variation

To prioritize genes associated with the regulation of blood pressure in the rat, we carried out a genome-wide correlation analysis between eleven BP parameters and gene transcript abundances in seven tissues (QTT analysis). Figure 4 gives the number and distribution of genes [cis-regulated, trans-regulated and not under genetic control (NGC transcripts)] that were most strongly associated with any BP phenotype across tissues (FDR, 5%). Overall, we found more than four-fold more NGC transcripts associated with any BP phenotype (ranging from 18.8–28.2% of transcripts across tissues) than either cis- (1.1–2.5%) or trans-eQTLs (2.2–6.5%). This suggests that variation in gene expression that is not under genetic control in the rat may either contribute to regulation of BP traits or might reflect changes secondary to BP (Supplementary material online, Table S6). In particular, using spectral and wavelet analysis we derived new BP parameters (Table 1), which significantly correlated with variation in gene expression across tissues similarly to that observed for standard BP indices (Figure 5). A large proportion of genes (37–59%, across tissues) were uniquely associated with the newly derived BP parameters (i.e. spectral and wavelet phenotypes), suggesting that these traits reflect different aspects of BP variation that was not captured by standard BP indices. Although we found more trans-eQTL genes associated with BP phenotypes overall, no differences were observed between the relative proportions of cis- and trans-eQTLs associated with standard and spectral/wavelet-derived BP traits (Figure 5).

3.4 Widespread conservation of eQTL genes between the rat and human hypertension

Genes that associated with cardiovascular traits in the rat, when conserved across species, can have similar functional consequences on related human phenotypes and disease. Here, we have performed a systematic investigation of all eQTL genes and transcripts analysed in seven tissues in the context of BP variation in the rat. We then carried out comparative genomics analysis between rat and human genes, focusing on human genes that were associated with hypertension in human GWAS. To this aim we compiled and annotated all genetic associations from 15 publicly available human GWAS and meta-analyses of hypertension and BP (see Section 2 and Supplementary material online, Table S7). We first explored whether rat eQTL genes whose expression correlated with BP variation were conserved between the rat and human genes associated with hypertension by GWAS (Supplementary material online, Table S8). Out of 2292 rat genes that were orthologue to human genes identified by GWAS, 795 (35%) were associated with BP phenotypes in the rat (FDR <5%, Supplementary material online, Table S9). Five hundred and ten (64%) of these genes formed eQTLs in the rat and were regulated in cis (n = 51), trans (n = 342), or showed both cis- and trans-regulation across tissues (n = 117). Overall, we found about seven-fold more trans- than cis-eQTLs that correlated with BP variation in the rat and that in human were associated (marginally or significantly) with hypertension in GWAS. Functional annotation analysis revealed
Figure 3  Trans-eQTL clusters detected across rat tissues. For each tissue the number of trans-eQTLs is reported as a fraction of the total number of eQTLs across the rat genome (chromosomes). A trans-eQTL cluster comprises at least 10 unique genes regulated in trans by a single SNP. The clusters that replicate across tissues are indicated by their names and are detailed in Supplementary material online, Table S3. Star symbols indicate the sixteen trans-eQTL cluster loci that harboured putative ‘master regulator’ genes (Table 3), which in humans were associated with hypertension by GWAS.
Figure 4  Rat genes associated with blood pressure phenotypes. For each tissue the number (and %) of transcripts that were significantly associated with each BP phenotype by QTT analysis is reported (FDR <5%). Eleven BP phenotypes were grouped as ‘conventional’ (diastolic blood pressure, DBP; mean blood pressure, MBP; peak-to-peak distance, PDIST; heart rate, HR; pulse pressure, PP; and systolic blood pressure, SBP), ‘spectral’ (maximum frequency, FREQ; total power, POWER, and maximum spectra, SPEC), and ‘wavelet’ (wave amplitude, WAVE and wave frequency, WFREQ).
Rat genes that associated with conventional and newly derived (spectral and wavelet) blood pressure traits. Left, for each tissue the Venn diagrams give the number of transcripts that significantly correlated with the "conventional" phenotypes (DBP, HR, MBP, PDIST, PP, and SBP) when compared with the "spectral" and "wavelet" phenotypes (FREQ, POWER, SPEC, WAVE, and WFREQ). The percentages are given with respect to the total number of transcripts analysed and are corrected for the number for probesets present in the Affymetrix microarrays. Right, number of cis- and trans-eQTLs that across tissues were associated with "conventional" and "spectral/wavelet" BP phenotypes, respectively. 'Common eQTLs' indicate eQTLs that were associated with both 'standard' and 'spectral/wavelet' BP-phenotypes, whereas 'unique eQTLs' indicate eQTLs that were specifically associated with either 'conventional' or 'spectral/wavelet' BP phenotypes.
that BP-associated trans-eQTLs were most highly enriched for ‘calcium signalling pathway’ \( (P = 9.6 \times 10^{-5}) \) and ‘ion channel’ related functions \( (P = 3.5 \times 10^{-5}) \), possibly reflecting the downstream effect of primary hypertension genes on pathophysiological pathways.

We then investigated to what extent rat gene sets (cis-, trans-regulated, and NGC-transcripts) were significantly enriched among the rat–human orthologous genes associated with hypertension by GWAS (Figure 1). In contrast with NGC transcripts, eQTLs were significantly over-represented in the rat–human orthologous genes (enrichment \( P < 0.017 \)), irrespective of whether they associated with BP traits in the rat or the tissue where they were expressed (Table 2). Specifically, both cis-eQTLs \( (14.4\%\text{, enrichment } P = 0.018) \) and trans-eQTLs \( (13.6\%\text{, enrichment } P = 0.021) \) were similarly over-represented in the human GWAS results more than what expected by chance. We also looked at eQTLs that overlapped within known BP-QTLs in the rat \( (n = 3713\text{ eQTLs; Supplementary material online, Table S3}) \) and found a significant over-representation for these genes in the human GWAS data (enrichment \( P = 2 \times 10^{-5} \)).

Although clusters of trans-eQTLs as a whole were not significantly over-represented in human GWAS, 22 individual clusters showed significant over-representation (enrichment \( P < 0.05 \)), two of these overlapped with known BP-QTLs in the rat and seven clusters contained at least 50% of genes significantly associated with BP traits in the rat (FDR<5%). The most important over-representation of rat genes in human GWAS results was observed for the putative ‘master regulators’ that underlie trans-eQTL cluster loci: 16 out of 56 ‘master regulator’ genes were found to be associated to hypertension in human GWAS, representing a significant enrichment over what expected by chance \( (28.6\%\text{; enrichment } P = 5 \times 10^{-4}) \). Three of these genes \( (Dpysl5, Bcat1, \text{ and } D3ZCG8\_RAT) \) were located at trans-eQTL clusters \( (aorta6\_25, skm4\_182, skm7\_90, \text{ Figure 3}) \) comprising genes that were over-represented among the human GWAS gene set (enrichment \( P < 0.05 \), Table 3). These analyses identified trans-eQTL clusters and their putative ‘master regulators’ that are conserved between rats and humans, and which are likely to play a role in human hypertension as primary and secondary determinants, respectively.

### 4. Discussion

In this study we have shown that genes associated with hypertension in human GWAS, when conserved with the rat, are likely to form both cis- and trans-acting eQTLs in multiple tissues. Our analysis of genetic regulation of gene expression and BP in the rat suggested that BP-associated genes are more likely to exert downstream effects on gene expression in trans. This emphasizes the role of secondary regulatory mechanisms in disease susceptibility, which are usually undetectable and unappreciated in human GWAS.

We used the HXB/BXH rat RI strains, a well-characterized model of hypertension, to assess the genetic regulation of gene expression in tissues important for the primary regulation of blood pressure or that possess an intrinsic capability to regulate the local blood flow (adrenal, kidney, and aorta). Likewise, we investigated genome-wide expression in other four tissues (liver, skeletal muscle, heart, and fat) that may respond to differences in blood pressure. Using a large data set of ~27 000 eQTLs we observed that, in addition to cis-regulated genes correlating with BP traits, trans-eQTLs can be associated with BP variation. Clustering of trans-acting eQTLs forming ‘hot spots’ \( ^{28} \) has been reported in several model organisms including yeast, \( ^{39} \) drosophila, \( ^{40} \) mice, \( ^{41} \) rat, \( ^{15,37} \) and suggested to be important in humans as well. \( ^{42,43} \) Although clusters of trans-regulated genes can harbour genes directly associated with several traits, including metabolic \( ^{44} \) and autoimmune phenotypes, \( ^{9} \) these have commonly smaller effect sizes than cis-regulated genes. \( ^{45} \) hence are more difficult to detect in model organisms and human GWAS. Here, taking advantage of powerful Bayesian eQTL mapping methods, \( ^{19} \) we identified a larger set of trans-eQTLs and trans-eQTL clusters, where many...
<table>
<thead>
<tr>
<th>Trans-eQTLs cluster (number of genes)</th>
<th>GO enrichment for the trans-eQTL cluster genes (P-value)</th>
<th>P-value of enrichment of the trans-eQTL cluster genes in human GWAS data</th>
<th>Rat gene underlying trans-eQTL cluster (within 2 kb window)</th>
<th>Significantly correlated blood pressure phenotype in the rat (FDR &lt;5%)</th>
<th>Human orthologue gene symbol</th>
<th>Human orthologue gene name</th>
<th>GWAS SNP</th>
<th>Gene Section</th>
<th>GWAS P-value</th>
<th>PMID for the GWAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>kid4_90 (n = 12)</td>
<td>—</td>
<td>0.07</td>
<td>Fam190a</td>
<td>—</td>
<td>FAM190A</td>
<td>Family with sequence similarity 190, member A</td>
<td>rs17018584</td>
<td>3′ UTR</td>
<td>2.45 × 10^{-11}</td>
<td>17554300</td>
</tr>
<tr>
<td>skm7_90 (n = 10)</td>
<td>—</td>
<td>0.04</td>
<td>DJZCG8_RAT</td>
<td>—</td>
<td>SAMD12</td>
<td>Sterile alpha motif domain containing 12</td>
<td>rs17829808</td>
<td>20′ 19′9SS</td>
<td>4.60 × 10^{-7}</td>
<td>17554300;17903302</td>
</tr>
<tr>
<td>skm4_182 (n = 25)</td>
<td>—</td>
<td>0.02</td>
<td>Bcat1</td>
<td>MBP, SBP; BCA1</td>
<td>BCA1</td>
<td>Branched chain amino-acid transaminase 1, cytosolic</td>
<td>rs7961152</td>
<td>Intron</td>
<td>7.39 × 10^{-4}</td>
<td>17554300</td>
</tr>
<tr>
<td>aorta19_12 (n = 10)</td>
<td>GO:0001890—placenta development (P = 7.5 × 10^{-4})</td>
<td>0.19</td>
<td>Large</td>
<td>LARGE</td>
<td>LARGK</td>
<td>Lysyl-glycosyltransferase</td>
<td>rs2277840</td>
<td>Intron</td>
<td>1.42 × 10^{-5}</td>
<td>17554300;21909115</td>
</tr>
<tr>
<td>skm7_87 (n = 14)</td>
<td>—</td>
<td>0.10</td>
<td>Trps1</td>
<td>MBD, DBP; TRPS1</td>
<td>TRPS1</td>
<td>Trichorhinophalangeal syndrome I</td>
<td>rs16887447</td>
<td>Intron</td>
<td>5.79 × 10^{-5}</td>
<td>17554300</td>
</tr>
<tr>
<td>fat1_154 (n = 10)</td>
<td>—</td>
<td>0.20</td>
<td>Oda4</td>
<td>—</td>
<td>ODZ4</td>
<td>Odd Oz/Ten-m homolog 4 (Drosophila)</td>
<td>rs648443</td>
<td>Intron</td>
<td>3.50 × 10^{-4}</td>
<td>17554300;21909115</td>
</tr>
<tr>
<td>lv8_1 (n = 10)</td>
<td>—</td>
<td>0.20</td>
<td>Vit</td>
<td>—</td>
<td>VIT</td>
<td>Vitrin</td>
<td>rs1468811</td>
<td>Intron</td>
<td>3.64 × 10^{-4}</td>
<td>17554300</td>
</tr>
<tr>
<td>aorta2_172 (n = 12)</td>
<td>—</td>
<td>0.27</td>
<td>Pdgfc</td>
<td>SPEC</td>
<td>PDGFC</td>
<td>Platelet-derived growth factor C</td>
<td>rs17230544</td>
<td>nonGene</td>
<td>3.83 × 10^{-4}</td>
<td>17463246</td>
</tr>
<tr>
<td>aorta2_234 (n = 153)</td>
<td>GO:0030029—actin filament-based process (P = 7.7 × 10^{-4})</td>
<td>0.38</td>
<td>Ppp3ca</td>
<td>—</td>
<td>PPP3CA</td>
<td>Protein phosphatase 3, catalytic subunit, alpha isozyme</td>
<td>rs2850341</td>
<td>Intron</td>
<td>3.96 × 10^{-4}</td>
<td>17554300</td>
</tr>
<tr>
<td>skm17_92 (n = 35)</td>
<td>—</td>
<td>0.08</td>
<td>Myo3a</td>
<td>—</td>
<td>MYO3A</td>
<td>Myosin III A</td>
<td>rs982082</td>
<td>Intron</td>
<td>4.91 × 10^{-4}</td>
<td>21909115</td>
</tr>
<tr>
<td>liver1_149 (n = 30)</td>
<td>GO:0005739—mitochondrion (P = 3.8 × 10^{-5})</td>
<td>0.24</td>
<td>Dlg2</td>
<td>—</td>
<td>Dlg2</td>
<td>Discs, large homolog 2 (Drosophila)</td>
<td>rs17146788</td>
<td>Intron</td>
<td>5.36 × 10^{-4}</td>
<td>17463246;17554300;21909115</td>
</tr>
<tr>
<td>lv8_28 (n = 31)</td>
<td>GO:0009057—macromolecule catabolic process (P = 2.6 × 10^{-5})</td>
<td>0.93</td>
<td>Opcml</td>
<td>—</td>
<td>OPCR</td>
<td>Opioid binding protein/cell adhesion molecule-like</td>
<td>rs11223077</td>
<td>Intron</td>
<td>5.50 × 10^{-4}</td>
<td>17554300;21909115</td>
</tr>
<tr>
<td>skm8_28 (n = 19)</td>
<td>—</td>
<td>0.20</td>
<td>Opcml</td>
<td>—</td>
<td>OPCR</td>
<td>Opioid binding protein/cell adhesion molecule-like</td>
<td>rs11223077</td>
<td>Intron</td>
<td>5.50 × 10^{-4}</td>
<td>17554300;21909115</td>
</tr>
<tr>
<td>aorta6_25 (n = 15)</td>
<td>—</td>
<td>0.03</td>
<td>DpySL5</td>
<td>—</td>
<td>DpySL5</td>
<td>Dihydropyrimidinase-like 5</td>
<td>rs827884</td>
<td>Intron</td>
<td>6.67 × 10^{-4}</td>
<td>17554300</td>
</tr>
<tr>
<td>lv3_161 (n = 10)</td>
<td>—</td>
<td>0.57</td>
<td>Cyp24a1</td>
<td>—</td>
<td>CYP24A1</td>
<td>Cytochrome P450, family 24, subfamily A, polypeptide 1</td>
<td>rs6022994</td>
<td>Intron</td>
<td>6.91 × 10^{-4}</td>
<td>17554300</td>
</tr>
<tr>
<td>aorta4_10 (n = 12)</td>
<td>GO:0006955—immune response (P = 4.4 × 10^{-5})</td>
<td>0.64</td>
<td>Magi2</td>
<td>—</td>
<td>MAGI2</td>
<td>Membrane-associated guanylate kinase, WW and PDZ domain containing 2</td>
<td>rs17150624</td>
<td>Intron</td>
<td>7.12 × 10^{-4}</td>
<td>17554300;21909115</td>
</tr>
</tbody>
</table>

*Clusters of trans-eQTLs were defined as a group of 10 or more genes mapping to the same locus (reported in Figure 3 and Supplementary material online, Table S3). Clusters detected in multiple tissues are highlighted with bold fonts.

Significance for gene ontology (GO) enrichment reported with FDR <5%.

Significance of rat trans-eQTLs cluster genes in the rat—human orthologous genes that were genetically associated with hypertension by GWAS (P < 0.001).

SNP from the GWAS which were reported to associated with hypertension.

Significance of the genetic association with hypertension reported in the GWAS.
clusters were replicated across tissues (Figure 3). Thirteen clusters (8%) were found at loci that overlap with known BP QTLs identified in SHR or SHR/BN strain combinations (Supplementary material online, Table S3), suggesting a role for these genes in blood pressure regulation in the rat.

To explore on a broader scale the contribution of eQTL genes to hypertension and identify candidate genes, we used radiotelemetric BP measurements and derived 11 parameters of BP variation, which were assessed by time series, spectral, and wavelet analyses (Table 1). Within each tissue, ~30% of analysed genes were associated with any BP parameter (Figure 4) and >8000 eQTLs, cis- and trans-acting, were associated with BP variation, identifying primary or secondary candidates for the regulation of blood pressure in the rat. Using wavelet and spectral analysis, we generated new, heritable traits for BP variation, which captured variation of cardiovascular functions and circadian BP that was not recapitulated by the ‘conventional’ BP parameters. This allowed for the identification of 2042 rat genes whose expression associated only with the new BP phenotypes (FREQ, POWER, SPEC, WAVE, and WFRFREQ) across tissues (Supplementary material online, Table S6).

We went on and investigated whether cis-, trans-eQTLs, trans-eQTL clusters and their putative ‘master regulators’ are conserved between rats and humans and associated with hypertension. Using human GWAS data, we tested whether rat gene sets were more likely to be over-represented in human genes associated with hypertension and blood pressure than by chance (Figure 1). Following Johnson and O’Donnell, we adopted a relaxed ‘discovery scan’ P-value threshold of $P < 0.001$ to create a sizeable set of human association results from 15 large GWAS studies, permitting rat gene sets to be formally tested for over-representation (see Section 2). This strategy allowed us to compile a large database of human hypertension genes ($n = 4145$; Supplementary material online, Table S7), and the choice of a relaxed ‘discovery scan’ P-value threshold enabled an even representation of association results from all studies, including those that released complete GWAS data (i.e. results not filtered for genome-wide significance, $P = 10^{-8}$). In addition to highly significant (and replicated) GWAS hits, we explicitly aimed to use the genetic association signal of small effect (i.e. detected with low significance, $P < 0.001$) by GWAS, which can shed light on trans- regulatory mechanisms underlying disease, as previously demonstrated.

Although we detected substantially more NGC transcripts (more than four-fold) associated with any BP phenotype in the rat than either cis- or trans-eQTLs, we found a significant over-representation of genes in the human GWAS results only for genes that in the rat formed eQTLs (Table 2). This observation is in keeping with recent data on annotation of human eQTL from lymphoblastoid cell lines that are enriched for trait-associated SNPs in GWAS studies.

Intriguingly, the group of cis-regulated genes that associated with BP parameters in the rat (when considered altogether) was less over-represented within human GWAS hits than BP-associated trans-eQTLs. Similar over-representation of trans-eQTLs has been reported for other complex traits, as shown in recent studies of type 2 diabetes in humans, which revealed marked excess of trans-acting eQTLs within the top association signals. The 2859 trans-eQTLs that associated with BP in the rat might be representative of ‘pathogenic pathways’ that can also play a role in human hypertension.

Functional annotation of the BP-associated trans-eQTLs that were conserved across species showed the most significant enrichment for ‘calcium signalling pathway’ genes ($P = 9.6 \times 10^{-7}$), reflecting an over-representation for ion transporters, sodium/potassium voltage channels, and glutamate receptors in this gene set. This observation is in keeping with several genetic, pathophysiological, and pharmacological studies providing evidence that ion channel functions are important in regulating BP levels, and suggest a role for conserved trans-regulated rat genes as secondary determinants of BP and hypertension.

Although strong enrichment was observed for the set of eQTLs that overlapped within known BP-QTLs in the rat (enrichment $P = 2 \times 10^{-4}$), the most suggestive result was found for the set of putative ‘master regulators’. Sixteen genes (out of 56) that were located at the trans-cluster loci were conserved to humans and detected in GWAS, representing a significant enrichment of what expected by chance (enrichment $P = 5 \times 10^{-4}$, Table 3). Since the putative ‘master regulator’ genes might regulate (in trans) clusters or networks of genes in the rat, these are likely to be primary (up-stream) regulators of pathways for blood pressure and hypertension. These ‘master regulator’ genes in humans were associated with hypertension with different levels of significance (GWAS $P$-value range from $10^{-4}$ to $10^{-11}$), and the strongest association signal for the FAM190A gene (GWAS $P$-value $= 2.4 \times 10^{-4}$). In the rat, Fam190a underlies a trans-eQTL cluster in the kidney tissue, which is moderately but not significantly enriched for human GWAS signals (enrichment $P = 0.07$, Table 3).

Among the trans-regulated genes of this cluster, potassium voltage-gated channel, shaker-related subfamily, member 5 (KCNA5) has been implicated with pulmonary hypertension, and atrial fibrillation. The second strongest GWAS signal within the set of putative ‘master regulators’ in humans was observed for sterile alpha motif domain containing 12 (SAMD12) (GWAS $P$-value $= 4.6 \times 10^{-3}$), which in the rat lies at a trans-eQTL cluster that is enriched for genetic associations in human hypertension (enrichment $P = 0.04$, Table 3). Another putative ‘master regulator’ rat gene that exerts a function in human cardiovascular diseases is branched chain amino-acid transaminase 1, cytosolic or BCAT1. BCAT1 was previously reported to be associated with hypertension (GWAS $P$-value $= 7.4 \times 10^{-3}$), as well as other cardiovascular traits, including salt sensitivity, and resting heart rate. Rat Bcat1 resides at the locus regulating trans 25 genes in skeletal muscle tissue, where 5 of these (Fat1, Miff1, A2bp1, Tyr, and Anxa5) were associated with hypertension in human GWAS, and Bcat1 skeletal muscle expression significantly correlated with MBP and SBP in the rat.

Genetically regulated changes in gene expression can determine phenotypic variation and susceptibility to disease. Cis-acting genetic control of gene expression has been widely investigated since it can be detected with relatively higher confidence than trans-acting control, and cis-acting eQTL genes can have large effects on whole body phenotypes and disease. In contrast, the specific role of trans-regulated genes in the context of disease susceptibility has been explored to a less extent. Our data support the hypothesis that secondary regulatory mechanisms of gene expression (trans-eQTLs) are largely conserved between rats and human disease. Extensive comparative genomics analysis of GWAS data showed that human hypertension genes, when conserved with the rat, are enriched for trans-eQTLs that associated with BP variation across multiple tissues. This most likely reflects the downstream effect of primary disease-associated genes on secondary regulatory mechanisms important for the pathogenesis of hypertension. These findings can help to elucidate the functional context and pathways mediating the effect of primary susceptibility genes identified in human GWAS.
of hypertension, highlighting the important contribution of trans-regulatory mechanisms for disease pathogenesis.

Supplementary material

Supplementary material is available at Cardiovascular Research online.

Acknowledgements

We thank the anonymous Reviewers for their critical and constructive comments on the manuscript.

Conflict of interest: none declared.

Funding

We acknowledge funding from European Community’s Seventh Framework Programme (FP7/2007-2013) under grant agreement no. HEALTH-F4-2010-241504 (EURATRANS) (E.P., T.J.A., S.A.C); the Medical Research Council (E.P., T.J.A., S.A.C., L.B. (G1002319)); the Wellcome Trust (S.R.L., L.B.); the Imperial BHF Centre of Excellence (T.J.A.); the Leduq Transatlantic Network of Excellence (T.J.A., S.A.C.). M.P. was supported by grants 7TE0067, LH11049 and grant LL1204 (within the ERC CZ program) from the Ministry of Education, Youth, and Sports of the Czech Republic.

References

2. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661–678.
46. Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. PLoS Genet 2010;6:e1000888.


