Transcriptional network analysis in muscle reveals AP-1 as a partner of PGC-1α in the regulation of the hypoxic gene program

Baresic, Mario and Salatino, Silvia and Kupr, Barbara and van Nimwegen, Erik and Handschin, Christoph

Posted at edoc, University of Basel
Official URL: http://edoc.unibas.ch/dok/A6288881

Originally published as:
Transcriptional network analysis in muscle reveals AP-1 as a partner of PGC-1α in the regulation of the hypoxic gene program

Mario Baresic¹,², Silvia Salatino¹,²,³, Barbara Kupr¹, Erik van Nimwegen²,³,* and Christoph Handschin¹,*

¹Focal Area Growth & Development and ²Focal Area Computational & Systems Biology, Biozentrum, University of Basel, Basel 4056, Switzerland
³Swiss Institute of Bioinformatics, Basel 4056, Switzerland

*These authors contributed equally to this manuscript


Copyright © the American Society for Microbiology; Molecular and Cellular Biology
Reprint request

Unfortunately, due to copyright-related issues, we are not able to post the post-print pdf version of our manuscript - in some cases, not even any version of our manuscript. Thus, if you would like to request a post-production, publisher pdf reprint, please click send an email with the request to christoph-dot-handschin_at_unibas-dot-ch (see http://www.biozentrum.unibas.ch/handschin).

Information about the Open Access policy of different publishers/journals can be found on the SHERPA/ROMEO webpage: http://www.sherpa.ac.uk/romeo/

Reprint Anfragen


Informationen zur Open Access Handhabung verschiedener Verlage/Journals sind auf der SHERPA/ROMEO Webpage verfügbar: http://www.sherpa.ac.uk/romeo/
Transcriptional network analysis in muscle reveals AP-1 as a partner of PGC-1α in the regulation of the hypoxic gene program

Mario Baresic¹,#, Silvia Salatino¹,²,³,#, Barbara Kupr¹, Erik van Nimwegen²,³,* and Christoph Handschin¹,*

¹Focal Area Growth & Development and ²Focal Area Computational & Systems Biology, Biozentrum, University of Basel, Basel 4056, Switzerland

³Swiss Institute of Bioinformatics, Basel 4056, Switzerland

#These authors contributed equally to this manuscript

*Correspondence to: Christoph Handschin, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland, email: christoph.handschin@unibas.ch +41 (0)61 267 23 78 or Erik van Nimwegen, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland, email: erik.vannimwegen@unibas.ch +41 (0) 61 267 15 76.

**Short title:** Transcriptional network regulation by PGC-1α

**Keywords:** skeletal muscle; PGC-1α; coregulators; transcriptional regulation; transcription factor binding site prediction; transcriptional networks; motif activity response analysis; principal component analysis; exercise
ABSTRACT

Skeletal muscle tissue shows an extraordinary cellular plasticity, but the underlying molecular mechanisms are still poorly understood. Here we use a combination of experimental and computational approaches to unravel the complex transcriptional network of muscle cell plasticity centered on the peroxisome proliferator-activated receptor γ coactivator 1α (PGC-1α), a regulatory nexus in endurance training adaptation. By integrating data on genome-wide binding of PGC-1α and gene expression upon PGC-1α over-expression with comprehensive computational prediction of transcription factor binding sites (TFBSs), we uncover a hitherto underestimated number of transcription factor partners involved in mediating PGC-1α action. In particular, principal component analysis of TFBSs at PGC-1α binding regions predicts that, besides the well-known role of the estrogen-related receptor α (ERRα), the activator protein-1 complex (AP-1) plays a major role in regulating the PGC-1α-controlled gene program of hypoxia response. Our findings thus reveal the complex transcriptional network of muscle cell plasticity controlled by PGC-1α.
INTRODUCTION

A sedentary life style can lead to an imbalance between energy intake and expenditure and favors the development of a number of chronic diseases like obesity and type 2 diabetes. Regular exercise on the other hand is an effective way to reduce the risk for these lifestyle-related pathologies (1). The health benefits of exercise are at least in part induced by changes in skeletal muscle tissue. Muscle cells exhibit a high plasticity and thus a remarkably complex adaptation to increased contractile activity. For example, endurance training induces mitochondrial biogenesis, increases capillary density and improves insulin sensitivity (1, 2). To achieve such a complex plastic response, a number of different signaling pathways are activated in an exercising muscle, for example p38 MAPK-mediated protein phosphorylation events, increased intracellular calcium levels or the activation of the metabolic sensors AMP-dependent protein kinase (AMPK) and sirtuin-1 (SIRT1) (3). While the temporal coordination of the numerous inputs is not clear, all of the major signaling pathways converge on the peroxisome proliferator-activated receptor γ coactivator 1α (PGC-1α) to either induce Ppargc1a gene expression, promote post-translational modifications of the PGC-1α protein, or by doing both (4, 5). Upon activation, PGC-1α mediates the muscular adaptations to endurance exercise by coactivating various transcription factors (TFs) involved in the regulation of diverse biological programs such as mitochondrial biogenesis, angiogenesis, ROS detoxification or glucose uptake (3). Accordingly, transgenic expression of PGC-1α in mouse skeletal muscle at physiological levels not only induces mitochondrial biogenesis but also drives a fiber type conversion towards a more oxidative, slow-twitch phenotype (6) while muscle-specific Ppargc1a knockout animals exhibit several symptoms of pathological inactivity (7, 8).

Coregulators are part of multicomponent regulatory protein complexes that are well suited to translate external stimuli into changes in promoter and enhancer activities by combining various enzymatic activities to modulate histones and chromatin structure, and recruit other TFs (9). Thus, dynamic assembly of distinct coregulator complexes enables the integration of many different signaling pathways leading to a coordinated and specific regulation of entire biological programs by multiple TFs (10, 11). For example, PGC-1α not only recruits histone acetylases (12), the TRAP/DRIP/Mediator (13) as well as the SWI/SNF protein complexes (14), but also binds to and coactivates a myriad of different transcription factors, even
though a systematic inventory of TF binding partners has not been compiled yet (15). Thus, the specific control exerted by the PGC-1α-dependent transcriptional network might provide an explanation for the dynamic and coordinated muscle adaptation to exercise. Since PGC-1α in skeletal muscle not only confers a trained phenotype, but also ameliorates several different muscle diseases (16), the unraveling of the PGC-1α-controlled transcriptional network in skeletal muscle would be of great interest to identify putative therapeutic targets within this pathway.

Therefore, we aimed at obtaining a global picture of the co-regulatory activity of PGC-1α in skeletal muscle cells. More precisely, by combining data on the genome-wide binding locations of PGC-1α and the gene expression profiles in response to PGC-1α over-expression with comprehensive computational prediction of transcription factor binding site (TFBS) occurrence, we sought to unveil the biological processes that are regulated by PGC-1α, to identify the transcription factors that partner with PGC-1α, and to determine the mechanistic details of PGC-1α-regulated transcription. We not only mapped the locations on the DNA where PGC-1α was bound, but also delineated the target genes whose expression is either directly or indirectly affected by PGC-1α and identified novel putative transcription factor partners that mediated PGC-1α’s action. In particular, our results strongly suggest that the activator protein-1 (AP-1) complex is a major regulatory partner of PGC-1α, with AP-1 and PGC-1α together regulating the hypoxic response gene program in muscle cells in vitro and in vivo.
MATERIALS AND METHODS

Cell culture and siRNA transfection

C2C12 cells were grown in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 100 Units/ml penicillin and 100µg/ml streptomycin. To obtain myotubes, the C2C12 myoblasts were allowed to reach 90% confluence and the medium was changed to DMEM supplemented with 2% horse serum (differentiation medium) for 72 hours.

The siRNAs for the knockdown of NFE2L2, FOS, JUN, ATF3, NFYC, ZFP143, GTF2I, the non-targeting siRNA pool and the DharmaFECT1 transfection reagent were purchased from Dharmacon (Fisher Scientific) and the siRNA transfection was performed according to the Thermo Scientific DharmaFECT Transfection Reagents siRNA Transfection Protocol. Briefly, after three days of differentiation, the respective siRNAs (50nM final concentration) was added to the medium. 24h after siRNA transfection, the cells were infected with either the PGC-1α or GFP adenovirus. Then, 48h after adenoviral infection, the cells were harvested.

Differentiated C2C12 cells were infected with adenoviral (AV) shERRα (kindly provided by Dr. Anastasia Kralli, Scripps Research Institute, La Jolla, CA, USA) to knockdown and inactivate ERRα or shGFP as a control. The infected cells were kept in culture for 4 days. Afterwards, cells were infected with the AV-flag-PGC-1α or AV-GFP and kept in culture two additional days. As a supplement to the previously infected AV shERRα cells, 2µM of the ERRα inverse agonist XCT-790 were added. To the remaining cells, 0.02% DMSO as a vehicle were added to the differentiated medium. All the experiments have been performed in biological triplicates. For RNA isolation, TRIzol® was used according to the TRIzol® reagent RNA isolation protocol (Invitrogen). Three conditions were used for further analysis: AV-shGFP + AV-GFP + vehicle, AV-shGFP + AV–flag-PGC-1α + vehicle, AV-shERRα + AV-flag-PGC-1α + 2µM XCT-790.

ChIP and ChIP Sequencing

ChIP was performed according to the Agilent Mammalian ChIP-on-chip Protocol version 10.0. For each immunoprecipitation, approximately 1x10⁸ C2C12 cells were differentiated into myotubes and infected with AV-flag-PGC-1α. For cross-linking protein complexes to DNA
binding elements, the cells were incubated in a 1% formaldehyde solution for 10 minutes, followed by the addition of glycine to a final concentration of 125 mM to quench the effect of the formaldehyde. The cells were rinsed in 1xPBS, harvested in ice-cold 1xPBS using a silicone scraper and pelleted by centrifugation. The pelleted cells were either used immediately or flash frozen and stored for later. The cells were then lysed at 4°C using two lysis buffers containing 0.5% NP-40/0.25% Triton X-100 and 0.1% Na-deoxycholate/0.5% N-lauroylsarcosine, respectively. The chromatin was then sheared by sonication to obtain DNA fragments of about 100-600 bp in length. 50 μl of the sonicated lysate was saved as input DNA. The immunoprecipitation was performed overnight at 4°C using magnetic beads (Dynabeads® Protein G, Invitrogen), which were previously coated with monoclonal antibodies like the monoclonal ANTI-FLAG® M2 Antibody, Sigma for the ChIP of PGC-1α or with the monoclonal anti-c-Fos (9F6) rabbit antibody #2250, Cell Signaling for the ChIP of FOS. The beads carrying the precipitate were washed five times for the c-Fos antibody and six times for the flag antibody with RIPA buffer and once with TE that contained 50 mM NaCl to eliminate unspecific binding of DNA to the beads. For elution, the beads were resuspended in elution buffer containing 1% SDS, placed in 65°C water bath for 15 minutes and vortexed every 2 minutes. To reverse the cross-links, the samples were incubated at 65°C overnight. The following day, the RNA and the cellular proteins were digested using RNase A and proteinase K. The DNA was precipitated and the success of the chromatin immunoprecipitation was validated by semiquantitative real-time PCR. The ChIP experiments were performed in triplicates. The ChIP of PGC-1α was further used for sequencing. The ChIP-Seq experiment of over-expressed PGC-1α in C2C12 cells was performed in biological duplicates. At the joint Quantitative Genomics core facility of the University of Basel and the Department of Biosystems Science and Engineering (D-BSSE) of the ETH Zurich in Basel, DNA libraries were prepared using the standard Illumina ChIP-Seq protocol, as described by the manufacturer, and the immunoprecipitated samples sequenced on the Genome Analyzer II. In order to keep only high quality data, the sequenced reads were filtered based on the quality score of each read and its alignments. Read were retained when Phred score >= 20, read length >= 25 bps and number of wrongly called nucleotides (Ns) <= 2. Those reads that passed the filter, (6’711’717 for the first immunoprecipitated sample (IP), 36’580’431 for the second IP, 17’899’074 for the first Whole Cell Extract (WCE), and 35’525’221 for the second WCE), were aligned to the mouse
genome, UCSC mm9 assembly, using Bowtie version 0.12.7 (17) using parameters --best --strata -a --m 100. The number of aligned reads equaled 5’699’648 for the first IP sample, 16’053’370 for the first WCE, 21’448’059 for the second IP, and 32’244’584 for the second WCE.

Identification of bound regions

To identify regions that were significantly enriched in the ChIP, we passed a 200 bps long sliding window along the genome, sliding by 25 bps between consecutive windows, and estimated the fraction of all ChIP reads $f_{IP}$ that fall within the window, as well as the fraction $f_{WCE}$ of reads from the whole cell extract that fall in the same window (which we estimate from a 2000 bps long window centered on the same genomic location). A Z-score quantifying the enrichment in the ChIP of each window was computed as:

$$Z = \frac{f_{IP} - f_{WCE}}{\sqrt{\sigma^2_{IP} + \sigma^2_{WCE}}}$$

where $\sigma^2_{IP}$ and $\sigma^2_{WCE}$ are the variances of the IP and WCE read frequencies, which are given by:

$$\sigma^2_{IP} = \frac{f_{IP}^* (1 - f_{IP})}{N_{IP}} \text{ and } \sigma^2_{WCE} = \frac{f_{WCE}^* (1 - f_{WCE})}{N_{WCE}}$$

respectively.

The enrichments were reproducible across biological replicates. Using only the first sequencing dataset, we called peaks at a Z cutoff of 4.5; we then compared these with the Z scores from the corresponding regions of the second dataset and the Pearson correlation coefficient was found to be 0.778. Similarly, we called peaks at a Z cutoff of 4.5 using only the second sequencing dataset; when we compared these peaks with the Z scores of the corresponding regions from the first dataset, the Pearson correlation coefficient was found to be 0.782.

To obtain a final set of binding peaks, we combined the reads from the two biological replicates computing the Z score of each window was computed as:
We conservatively considered all windows with a Z-score larger than 4.5 as were considered significantly enriched (False Discovery Rate 0.6%). The final binding peaks were obtained by merging consecutive windows that all passed the cut-off and by considering the “peak” to correspond to the top scoring window, i.e. corresponding to the summit of the ChIP-Seq signal. To determine the PGC-1α distribution genome-wide, peaks were annotated according to their closest Mus musculus RefSeq transcripts. We defined peaks as: “Intronic” (peak center lying inside an intron); “Exonic” (peak center lying inside an exon); “Upstream of TSS” (peak center lying within -10 to 0 kb from the closest TSS); “Downstream of TES” (peak center lying within 0 to 10 kb from the closest TES); “Intergenic” (peak center located farther than 10 kb from the nearest transcript). Moreover, we computed the ratio between observed and expected peak location distributions, obtained by generating 100 peak sets composed of 7512 random peaks each.

Motif finding and TFBSs over-representation

The binding peak regions were aligned to orthologous regions from other 6 mammalian species – human (hg18), rhesus macaque (rheMac2), dog (canFam2), horse (equCab1), cow (bosTau3) and opossum (monDom4) – using T-Coffee (18). A collection of 190 mammalian regulatory motifs (position weight matrices or WMs) representing the binding specificities of approximate 350 mouse TFs (in many cases, sequence specificities of multiple closely-related TFs were represented with the same WM) were downloaded from the SwissRegulon website (19). TFBSs for all known motifs were predicted using the MotEvo algorithm (20) on the alignments of all the 7512 peak sequences. Only binding sites with a posterior probability $\geq 0.1$ were considered for the further steps of the analysis. In order to create a background set of regions to assess the overrepresentation of binding sites within our regions, we created randomized alignments by shuffling the multiple alignment columns, maintaining both the gap patterns and the conservation patterns of the original alignments. TFBSs were predicted on the shuffled alignments using the same MotEvo settings as for the original peak alignments. Over-representation of motifs in the PGC-1α binding peaks was calculated by comparing total predicted TFBS occurrence within binding peaks with the predicted TFBS
occurrence in the shuffled alignments. We evaluated the enrichment of TFBSs for each motif \( x \) by collecting the sum \( n_x \) of the posterior probabilities of its predicted sites in the peak alignments as well as the corresponding sum \( n'_x \) in the shuffled alignments, and computed a Z-score:

\[
Z = \frac{f_x - f'_x}{\sqrt{\frac{f_x (1 - f_x)}{L_x} + \frac{f'_x (1 - f'_x)}{L'_x}}}
\]

where \( L_x \) and \( L'_x \) are the total lengths of the original and shuffled alignments, respectively, while \( f_x \) and \( f'_x \) are given by the equations:

\[
n_x \cdot l_x = f_x \cdot L_x \quad \text{and} \quad n'_x \cdot l_x = f'_x \cdot L'_x
\]

with \( l_x \) the length of motif \( x \).

**Principal Component Analysis of TFBS occurrence in binding peaks**

The input matrix \( N \) for the Principal Component Analysis (PCA) contained the total number of predicted binding sites \( N_{pm} \) in each of the 7512 binding peaks \( p \) (rows) for each of the 190 mammalian regulatory motifs \( m \) (columns). After mean centering the columns of this matrix, \( \bar{N}_{pm} = N_{pm} - \langle N_m \rangle \), i.e. subtracting the average site count for each motif, Singular Value Decomposition (SVD) was used to factorize this matrix: \( \bar{N} = U \cdot S \cdot V^T \), where \( U \) is a \( P \times M \) matrix whose columns are the left singular vectors of \( \bar{N} \); \( S \) is a \( M \times M \) diagonal matrix containing the singular values, and \( V^T \) (the transpose of \( V \)) is an \( M \times M \) matrix whose rows are the right singular vectors, with \( P \) the number of peaks, and \( M \) the number of motifs. The SVD was performed using the “svd” package of the “R” programming language.

**Gene expression arrays**

Whole-gene expression after 48 hours of transfection with adenovirus was measured in C2C12 cells with Affymetrix GeneChip® Mouse Gene 1.0 ST microarrays at the Life Science Training core facility of the University of Basel. Raw probe intensities were corrected for background and unspecific binding using the Bioconductor package “affy” (21).
Subsequently, probes were classified as expressed or non-expressed by using the “Mclust” R package (22) and, after removal of non-expressed probes, the intensity values were quantile normalized across all samples. Using mapping of the probes to the UCSC collection of mouse mRNAs, probes were then associated to a comprehensive collection of mouse promoters available from the SwissRegulon database (19). The log2 expression level of a given promoter was calculated as the weighted average of the expression levels of all probes associated to it. Log2 expression levels were then compared between over-expressed PGC-1α and the control GFP sample; for each promoter, the change in expression level across the two conditions was measured by log2 fold change (log2FC), computed as the difference between the mean of the log2 values in PGC-1α and the mean of the log2 values in GFP. The significance of the expression change was assessed by a Z score, which was computed as:

\[
Z = \frac{\bar{E}_{\text{PGC1}\alpha} - \bar{E}_{\text{GFP}}}{\sqrt{\frac{\sigma^2_{\text{PGC1}\alpha}}{n} + \frac{\sigma^2_{\text{GFP}}}{n}}}
\]

where \( n = 3 \) was the number of replicate samples, \( \bar{E}_{\text{PGC1}\alpha} \) is the mean log2 expression across the PGC-1α samples, \( \bar{E}_{\text{GFP}} \) is the mean log2 expression across the GFP samples, and \( \sigma^2_{\text{PGC1}\alpha} \) and \( \sigma^2_{\text{GFP}} \) are the variances of log2 expression levels across the replicates for the PGC-1α and control samples, respectively. Promoters were considered significantly up-regulated when log2FC \( \geq 1 \) and \( Z \geq 3 \), and significantly down-regulated when log2FC \( \leq -1 \) and \( Z \leq -3 \).

Peaks were assigned to promoters by proximity. To assign each peak to a promoter, we calculated the distance from the center of the peak to the center of neighboring promoters; whenever the peak was closer than 10 kb from at least one promoter, it was assigned to the nearest promoter and, thus, to its associated gene.

**Gene Ontology enrichment analysis**

Gene IDs were extracted from differentially regulated promoters and divided in four groups: up-regulated promoters with an assigned binding peak, up-regulated promoters without an assigned binding peak, down-regulated promoters with an assigned peak, and down-regulated promoters without an assigned peak. These four gene sets were used as input for
the functional analysis tool FatiGO (23) to identify significantly over-represented Gene Ontology (GO) categories compared to all Mus musculus genes. Only GO terms having an FDR-adjusted p-value <= 0.05 were considered significant.

**Motif activity at direct and indirect targets of PGC-1α**

To integrate the information from the PGC-1α binding peaks, we extended MARA (24) to model the direct and indirect regulatory effects of PGC-1α. Given the input expression data and the computationally predicted binding sites, MARA infers, for each of 190 regulatory motifs $m$, the activity $A_{ms}$ of the motif in each sample $s$ when the motif occurs *outside* of a region of PGC-1α, and the activities $A^*_{ms}$ of the motifs when they occur *within* a PGC-1α binding peak. That is, changes in the motif activities $A_{ms}$ upon over-expression of PGC-1α indicate indirect regulatory effects of PGC-1α on each motif $m$, whereas changes in the motif activities $A^*_{ms}$ reflect direct regulatory effects of PGC-1α as mediated by each motif $m$. For each promoter $p$ that was not associated with any PGC-1α binding peak (which we denote indirect targets), we modeled its log-expression in sample $s$, $e_{ps}$, in terms of the predicted number of TFBSs $N_{pm}$ that occur in the proximal promoter region (running from -500 to +500 relative to TSS) for each regulatory motif $m$. That is, MARA assumes the linear model:

$$e_{ps} = c_p + \tilde{c}_s + \sum_m N_{pm}A_{ms}$$

where $c_p$ is the basal expression of promoter $p$, $\tilde{c}_s$ is a sample-dependent normalization constant, and $A_{ms}$ is the regulatory activity of motif $m$ in sample $s$, which is inferred by the model. Formally, $A_{ms}$ quantifies amount by which the expression of promoter $p$ in sample $s$ would be reduced if a binding site for motif $m$ were to be deleted from the promoter.

For each “direct target” promoter $p$ that has an associated PGC-1α binding peak, which we defined as promoters with a peak within 1 kb or with a peak within 100 kb that was highly conserved according to PhastCons score of the region (25), we model its expression in terms of the predicted TFBSs in the binding peak, i.e.:

$$e_{ps} = c_p + \tilde{c}_s + \sum_m N_{pm}^*A_{ms}^*$$
where $N_p^*$ is the number of predicted TFBSs for motif $m$ in the peak associated with promoter $p$, and $A_{ms}^*$ is the motif activity of regulator $m$ in sample $s$ when this motif occurs in the context of PGC-1α binding. That is, the inferred motif activities $A_{ms}$ quantify the activities of regulatory motifs when they occur independent of PGC-1α binding, and the motif activities $A'_{ms}$ quantify the activities of motifs when they occur in a PGC-1α binding peak, i.e. the latter activities reflect direct effects of a PGC-1α while the former reflect indirect effects.

MARA predicts activities for 190 different mammalian regulatory motifs, associated with roughly 350 mouse TFs. Besides motif activities MARA also calculates error-bars $\delta_{ms}$ for each motif $m$ in each sample $s$. Using these, MARA calculates, for each motif $m$, an overall significance measure for the variation in motif activities across the samples analogous to a $z$-statistic:

$$z_m = \sqrt{\frac{1}{S} \sum_{s=1}^{S} \left( \frac{A_{ms}}{\delta_{ms}} \right)^2}$$

For each motif we calculate both a $z$-score $z_m$ associated with its indirect activity changes, and a $z$-score $z_m^*$ associated with its direct activity changes. MARA also ranks the confidence on predicted target promoters of each motif by a Bayesian procedure that quantifies the contribution of that factor to explaining the promoter’s expression variation by a Chi-squared value (for details, see (24)). The parameters used for motif stratification were: (i) the Z score $z_m^*$ for direct activity changes, (ii) the Z score $z_m$ for indirect motif activity changes, (iii) the Z score $\bar{z}_m^*$ for direct motif activity changes, computed by averaging the sample replicates and (iv) the Z score $\bar{z}_m$ for indirect motif activity changes, computed by averaging the sample replicates. The latter two measures were used to show which direction the motif activity changes when over-expressing PGC-1α with respect to the control condition. All motifs $m$ for which either the direct or indirect motif activities were changing significantly ($z \geq 2$) were subsequently selected.

**De novo motif finding**
PhyloGibbs (26) was used to identify de novo motifs across the 200 top enriched PGC-1α peaks. The parameters used were -D 1 -z 1 -γ 200 -m 10, corresponding to searching on multiple alignments for a single motif of length 10 with a total of 200 sites. The resulting motif was scanned for similarity to the other known motifs from our dataset using STAMP (27), with settings: Pearson Correlation Coefficient for column comparison metric, Smith-Waterman for the alignment method, penalty of 0.5 and 0.25 for gap opening and gap extension, respectively.

**Real-time PCR and target gene validation**

Putative target genes of distinct transcription factor-PGC-1α combinations were chosen according to three criteria: first, positive transcriptional regulation by PGC-1α by more than 2 fold, second, presence of a PGC-1α binding peak within a 10 kb distance from the TSS and third, prediction of targeting by MARA with a positive Chi-squared score. The sequences of the primers used in real-time PCR experiments are depicted in Suppl. Table 1. Relative mRNA was quantified by qPCR on a StepOnePlus system (Applied Biosystems) using Power SYBR Green PCR Master Mix (Applied Biosystems).

The values are presented as the mean +/- SEM. A Student’s t-test was performed and a p-value < 0.05 was considered as significant (*p<0.05, **p<0.01, ***p<0.001).

**Animals**

Mice were housed in a conventional facility with a 12-h night/12-h day cycle with free access to chow diet pellet and water. For the experiments, 22-23 week-old skeletal muscle-specific HSA-PGC-1α knockout (MKO) male mice and 8 week-old PGC-1α muscle-specific transgenic (TG) male mice were used as previously described (6-8). All experiments were performed according to the criteria outlined for the care and use of laboratory animals and with approval of the veterinary office of the canton Basel and the Swiss authorities.

**Treadmill running**
Treadmill running was performed with the TG mice on the Columbus Instruments motorized treadmill with an electric shock grid. The mice were acclimatized to the treadmill and then let run till exhaustion. The running protocol is as follows: 10m/min for 5min with an increase by 2m/min every 5min until 26m/min and an inclination of 5 degrees. The speed of 26m/min was kept until exhaustion of the mice (7, 28, 29). Mice were killed and tissues were collected 3h after exercise.

**RNA isolation of muscle tissue**

Gastrocnemius and quadriceps were used to isolate RNA by TRIzol® according to the TRIzol® reagent RNA isolation protocol (Invitrogen).
RESULTS

Broad recruitment of PGC-1α to the mouse genome

PGC-1α-dependent gene transcription has been studied in many different experimental contexts. In isolation, gene expression arrays however are unable to distinguish direct from indirect targets, or to reveal the genomic sites where PGC-1α is recruited to enhancer and promoter elements, i.e. by coactivating TFs that directly bind to the DNA. Thus, we first performed chromatin immunoprecipitation followed by deep sequencing (ChIP-Seq) of PGC-1α in differentiated C2C12 mouse myotubes to identify the locations where PGC-1α is bound to the genome. To identify genomic regions that are significantly enriched in the ChIP, we slid a 200 bp window across the genome comparing the local ChIP read density with the read density from a background whole cell extract sample. We selected all regions with a Z-statistic larger than 4.5 as significantly enriched (FDR 0.6%, Suppl. Fig. S1A). Using this stringent cutoff, we identified 7512 binding regions for PGC-1α via interaction with a TF genome-wide, which include binding regions in the promoters of known PGC-1α target genes (Fig. 1A) such as medium-chain specific acyl-CoA dehydrogenase (Acadm) and cytochrome c (Cycs) (30, 31). The enrichment of immunoprecipitated DNA fragments from the ChIP-Seq was validated for these and other PGC-1α target genes by semiquantitative real-time PCR (Fig. 1B). In absolute terms, the distribution of the ChIP-Seq peaks revealed that PGC-1α is mostly recruited at distal sites from the assigned targets and, to a lesser extent, to proximal regions of the gene or within an intronic sequence (Fig. 1C). However, when compared to randomly selected DNA regions of equal size and number, PGC-1α binding peaks occur twice as often within 10 kb upstream of the transcription start site (TSS).

In parallel to the ChIP-Seq experiment, we furthermore analyzed gene expression patterns in differentiated muscle cells both in control condition and under PGC-1α over-expression. Using a reference set of mouse promoters (19) and associating microarray probes to promoters by mapping to known transcripts, we found 1566 promoters (corresponding to 984 genes) to be significantly up-regulated (log2 fold change >= 1; Z score >= 3) and 1165 promoters (corresponding to 727 genes) to be significantly down-regulated (log2 fold change <= -1; Z score <= -3). Thus, similar to previous reports, PGC-1α induced and repressed the transcription of almost the same number of genes, respectively, indicating that the
physiological function of PGC-1α includes both the activation and inhibition of substantial numbers of genes.

To combine the DNA binding results from the ChIP-Seq with the data of the gene expression arrays, we then assigned ChIP-Seq peaks to the closest promoter (and the associated gene) within a maximum distance of 10 kb. In this way, about 30% of all peaks (2295 of 7512) could be associated with a target promoter. Inversely, for about 35% of all significantly up-regulated genes (341 of 984), a PGC-1α binding peak is found within 10 kb of the promoter. Since some of the up-regulated promoters may be regulated by more distal peaks, this is only a lower bound on the fraction of genes that are directly regulated. In stark contrast, only about 5% of all repressed genes harbor one or more PGC-1α DNA recruitment peaks in their vicinity (36 of 727) opposed to 95% indirectly down-regulated PGC-1α target genes (691 genes) (Fig. 1D). Moreover, the distribution of the distances between PGC-1α peaks and their associated promoters revealed a tight cluster of 532 peaks close to promoter regions for up-regulated, direct PGC-1α target genes (Fig. 1E) whereas the distribution of the 43 peaks associated to down-regulated genes was much wider, raising the possibility that the association of peaks to transcriptionally repressed genes might be spurious (Fig. 1F). In summary, the strong enrichment of binding peaks near up-regulated genes and the almost complete absence of binding peaks near down-regulated genes suggest that direct regulation of transcription by PGC-1α is almost exclusively activating. We note that there is a large fraction of binding peaks (75%) that are associated to target genes that do not significantly alter their expression. These peaks may have been wrongly assigned, their functionality may be dependent on additional factors not active in these cells, or they may simply be spurious binding events that are not functional.

We next used this stratification of peaks and genes to study whether direct (i.e. with an associated binding peak) and indirect PGC-1α target genes exert different biological function and identified Gene Ontology (GO) terms that were over-represented in any of the four categories. First, we observed that the most significantly enriched functional categories for directly and indirectly up-regulated genes were those related to mitochondria, oxidative phosphorylation and energy production (Fig. 1G and Suppl. Fig. S1B). In contrast, GO analysis of indirectly down-regulated PGC-1α target genes revealed a high prevalence of terms related to inflammation and immune response (Fig. 1H and Suppl. Fig. S1C). Assuming that
the assignment of peaks to repressed genes is not spurious, the few directly repressed PGC-1α targets exhibit an enrichment in functions related to muscle contraction, in particular for genes that are linked to contractile and metabolic properties of glycolytic, fast-twitch muscle fibers (Fig. 1H and Suppl. Fig. S1D), as would be expected from the observed shift from glycolytic to oxidative fibers mediated by PGC-1α in muscle (6).

**Modeling the direct and indirect gene regulatory effects of PGC-1α**

As a next step, we rigorously modeled the effects of PGC-1α on its target genes in terms of the occurrence of TFBSs for a large collection of mammalian regulatory motifs. We previously introduced a general framework, called Motif Activity Response Analysis (MARA) (24), for modeling the gene expression profiles as a linear function of the TFBSs occurring in the promoters and unknown regulatory “activities” of each of the regulators. As detailed in the Methods, we here extended MARA to incorporate information from the PGC-1α ChIP-Seq data, with the aim of identifying which other TFs are involved in mediating both the direct and indirect regulatory effects of PGC-1α. Specifically, for all “direct target” promoters that were associated with a PGC-1α binding peak, we modeled the expression of the promoter in terms of the predicted TFBSs in the neighborhood of the binding peak, while for “indirect target” promoters we modeled the promoter’s expression in terms of the predicted TFBSs in the proximal promoter region, according to the conventional MARA approach (Fig. 2A and 2B).

First, further supporting our analysis above, direct target promoters were almost exclusively up-regulated and only in a few exceptional cases reached statistical significance for PGC-1α-repressed transcripts (Fig. 2C). Among the direct motif activities, the ESRRA position weight matrix was the top ranking motif with a Z score of 6.04 (Suppl. Fig. 2). The corresponding TF estrogen-related receptor α (ERRα), an orphan nuclear receptor, has been extensively studied as a central binding partner for PGC-1α in the regulation of mitochondrial gene expression (30-32). To stratify the different motifs according to their predicted function, we then divided all motifs into groups according to the behavior of both their direct and indirect activity changes. Strikingly, all motifs exhibited one of only four different motif activity patterns. First, 6 TFs (Suppl. Fig. S2) were predicted to positively regulate PGC-1α target genes only in the presence of PGC-1α (Fig. 2D). Second, we found 6 motifs (Suppl. Fig. S2)
with significantly up-regulated direct and indirect motif activities upon PGC-1α over-expression (Fig. 2E). To our surprise, ERRα was predicted to regulate PGC-1α target genes in this manner, even though in previous reports gene regulation by ERRα in the context of activated PGC-1α was suggested to be dependent on PGC-1α coactivation (30-32). Third, we found 13 motifs (Suppl. Fig. S2) that were predicted to regulate PGC-1α target genes, however only in the absence of PGC-1α (Fig. 2F). Fourth, there was a group of 28 motifs (Suppl. Fig. S2) that showed a significant decrease of indirect motif activity upon PGC-1α over-expression, but no significant change of their direct motif activity, including NFκB (Fig. 2G), a central regulator of inflammation which is indirectly repressed by PGC-1α (33). Intriguingly however, no motif was found that showed significant direct repression of target genes, reinforcing the hypothesis that PGC-1α-dependent gene repression is an indirect event.

Nuclear receptors and activator protein-1-like leucine zipper proteins are the main functional partners of PGC-1α in muscle cells

As a next step, we analyzed the occurrence of TF DNA-binding motifs in the PGC-1α peaks identified by ChIP-Seq. We first performed de novo motif prediction on the top 200 peaks, using PhyloGibbs (26). As shown in Figure 3A, the motif that PhyloGibbs identified matches significantly (E-value = 7.7834e-10, as calculated by STAMP (27)) the canonical ESRRα motif. In addition to the de novo prediction, we also used the same collection of 190 mammalian regulatory motifs used by MARA (19) to check which known TF DNA-binding motifs were significantly over-represented in the PGC-1α peaks relative to a set of background regions. Many of the most significantly enriched motifs represent variations of nuclear receptor binding sequences that are based on the “AG^T/C/CTCA” core hexamer and occur either alone or in direct, inverted or everted repeats with variable spacing (Fig. 3B). Of these, the most significantly enriched motif was ESRRα, which is present in ~20% of all peaks. Moreover, among all genes with at least one associated binding peak within 10Kb, ~28% are associated with a peak containing a predicted ERRα site. Interestingly, besides the nuclear receptor motifs, we also found the DNA-binding element of the insulator protein CCCTC-binding factor (CTCF), and a set of highly similar DNA elements sharing the FOS-JUN-like recognition
sequence “TGA^G/TCA” bound by the TFs BACH2, FOS, FOSB, FOSL1, JUN, JUNB, JUND, FOSL2, NFE2, and NFE2L2 among the top 15 motifs enriched in PGC-1α peaks (Fig. 3B).

The identity of the exact nuclear receptor binding partner that is bound at each peak is difficult to deduce from DNA-binding motifs, since considerable promiscuity exists between receptors and DNA-binding elements in different configurations of hexameric repeats (34). Moreover, non-nuclear receptor-like TFs are less well studied in the context of PGC-1α-controlled gene expression. Thus, to identify which regulatory motifs are most over-represented among peaks that do not contain nuclear receptor-like sites, we first manually grouped all of the motifs with a sequence logo very similar to that of ESRRA. Next, we discarded all peaks that had one or more predicted TFBSs for any of the motifs in this set. With the remaining 3856 DNA sequences (51.33% of the peaks), we then again assessed the over-representation of each of the 190 mammalian regulatory motifs. In this analysis, "TGA^G/TCA” recognition elements, hence FOS-JUN-like motifs, were the most significantly enriched among these peaks (Fig. 3C). This result suggests that PGC-1α peaks naturally fall into two classes: those containing ESRRA-like sites, and those containing sites for FOS-JUN-like motifs.

We then constructed a matrix $N$, whose elements $N_{pm}$ contain the number of predicted TFBSs for each motif $m$ in each peak region $p$. We then performed principal component analysis (PCA) on this site-count matrix to identify linear combinations of regulatory motifs that explain most of the variation in site-counts across the PGC-1α peaks. The first two components (out of 190 in total) clearly proved to be the most relevant ones, accounting for 10% and 9.6% of the total variation in our dataset, respectively (Fig. 3D). Figure 3E shows the projection of all motifs on these first two principal components, with the names of the motifs with the largest projection indicated in the figure. Whereas most motifs have projections close to zero along the first component, there is one group of motifs with strong negative projections (ESRRA, NR1H4, NR5A1, 2, NR6A1) and one group of motifs with strong positive projections (BACH2, FOS_FOS(B,L1)_JUN(B,D), FOSL2, NFE2, NFE2L1, NFE2L2). These two sets of sites correspond precisely to the two classes of motifs identified above, confirming that the most significant variation in TFBSs across PGC-1α peaks is caused by the occurrence of either ESRRA-like motifs, or FOS-JUN-like motifs. Most interestingly, these two clusters of motifs reflect structurally distinct classes of TFs; the negatively scoring
eigenmotifs are characterized by binding of nuclear receptor-type zinc finger domains, while the eigenmotifs with a positive score correspond to activator protein-1 (AP-1)-like leucine zipper domains.

The second principal component corresponds to the strength of the binding signal for these 10 motifs, as confirmed by the robust negative correlation ($r=-0.92$) between the TFBSs posterior sum per peak and the peak’s projection along the second principal component (Fig. 3F).

**Validation of top scoring motifs reveals novel functional partners of PGC-1α**

Our analysis identified a number of so-far uncharacterized TFs as potentially functional partners for PGC-1α-controlled gene expression in skeletal muscle cells. In order to experimentally validate some of these candidates, we sorted all TFs by a number of criteria including TFBS over-representation in binding peaks, MARA activity upon PGC-1α over-expression, and the expression pattern of the TFs themselves. Table 1 shows the top 15 ranked TFs according to this selection. As expected, the well-known PGC-1α partner ERRα was identified as the most important factor. For our validation experiments, we chose the next two motifs (FOS_FOS(B,L1)_JUN(B,D) and ZNF143, which is also known as ZFP143) as well as three motifs from further down the list of the top 15 motifs (GTF2I, NFE2L2 and NFYC).

FOS, the most up-regulated TF (log2 fold change = 1.78) among the TFs associated with the motif FOS_FOS(B,L1)_JUN(B,D), is a basic leucine zipper transcription factor known to heterodimerize with other leucine zipper proteins in order to form the AP-1 complex (35). The AP-1 complex furthermore contains JUN as well as ATF proteins. Thus, to dissect the function of the AP-1 protein complex, we also included JUN and ATF3, the most highly expressed isoforms of their respective protein families in muscle cells.

For each of these 7 TFs (ATF3, FOS, GTF2I, JUN, NFE2L2, NFYC and ZFP143), we selected a dozen target genes based on the Chi² score of the MARA prediction, presence of a PGC-1α binding peak with at least one predicted binding site for the factor of interest, and at least a 2-fold induction upon over-expression of PGC-1α. As summarized in Fig. 4 and Suppl. Fig. S3, siRNA-based knockdown of all TFs resulted in a robust reduction of the target mRNAs from -
40% to -75%. With the exception of NFYC and JUN, we found that the large majority of predicted target genes were down-regulated upon knockdown of the factor, confirming our predictions (Fig. 4). The most consistent effects were observed for FOS and ZFP143 (all targets down-regulated), followed by GTF2I (11 out of 12 down-regulated) and NFE2L2 and ATF3 (10 out of 12 down-regulated). Interestingly, distinct target genes of the AP-1 complex showed differential responsiveness to knockdown of the three AP-1 complex components FOS, JUN and ATF3 (Fig. 4B, Fig. 4C and Fig. 4D). Similarly, PGC-1α-mediated induction of a majority of the predicted target genes for NFE2L2 (Fig. 4E), ZFP143 (Fig. 4F) and GTF2I (Fig. 4G) was reduced upon knockdown of the respective TF when compared to the expression in cells with overexpressed PGC-1α and a scrambled siRNA control. Surprisingly, only 1 of the 11 predicted target genes for NFYC that have been chosen for validation was significantly repressed by siRNA-induced reduction of this TF (Fig. 4H), suggesting that other TFs may be involved in mediating the regulatory effects of the NFYC regulatory motif.

**Functional interaction between PGC-1α and different compositions of the AP-1 protein complex**

Our targeted validation strategy revealed that PGC-1α target genes predicted to be regulated by the FOS-JUN-like motif react in distinct manners to siRNA-mediated knockdown of individual components of the AP-1 transcription factor protein complex. For example, some genes only reacted to reduction of FOS (Fig. 5A), while others were responsive to the knockdown of two (Fig. 5B) or even all three AP-1 protein partners (Fig. 5C) that we have tested using the siRNA-based approach. To further dissect the responsiveness of PGC-1α target genes to different AP-1 protein complexes, we performed global gene expression arrays upon knockdown of each of the three TF components of the AP-1 complex. Fig. 5D depicts the number of genes that were induced by PGC-1α and that were, at the same time, down-regulated by the siRNA knockdown of any of the three AP-1 complex members. Amongst a total of 477 genes, 89% responded to FOS knockdown, 52% to ATF3 knockdown, and 31% to JUN knockdown. Moreover, while 37% of all targets responded exclusively to FOS, the fraction of targets responding exclusively to either JUN or ATF3 was at most 5%. This analysis shows that, whereas different target genes respond differently to the
knockdown of distinct AP-1 components, FOS is the dominant factor in determining AP-1 function in these conditions.

As shown in Fig. 3B, 341 genes were associated to a PGC-1α binding peak containing a predicted site for the FOS-JUN-like motif bound by the AP-1 complex. Of these genes, the expression of 55 was significantly induced by PGC-1α over-expression in muscle cells. In our siRNA-based validation experiment, we found that 47 out of these 55 PGC-1α-induced/AP-1 predicted targets were significantly down-regulated by knockdown of the AP-1 complex components and we called these genes “direct PGC-1α/AP-1 targets”. The remaining 430 genes out of 477 (Fig. 5D) were defined accordingly as “indirect PGC-1α/AP-1 targets” that lack a PGC-1α peak containing a FOS-JUN-like motif, but still are regulated by PGC-1α and the AP-1 protein components (Fig. 5E). To reveal whether these gene categories exert distinct functions, GO and KEGG enrichment analyses were performed. Surprisingly, the 47 direct PGC-1α/AP-1 target genes showed a distinct and significant over-representation of the terms “response to hypoxia” (GO ID: 0001666; adjusted p-value: 0.0247542) and “mTOR signaling pathway” (KEGG ID: mmu04150; adjusted p-value: 0.030674) that were absent in the GO analysis of the remaining PGC-1α/AP-1 targets (Fig. 5F). Recruitment of FOS to the same regulatory regions as PGC-1α in the direct AP-1/PGC-1α target genes was subsequently validated by ChIP (Fig. 5G). These results suggest that AP-1, when interacting with PGC-1α, drives a synergic effect of response to hypoxia; on the other hand, when AP-1 and PGC-1α act separately, and furthermore through downstream intermediate TFs, they regulate the expression of genes involved in mitochondrial organization and energy metabolism.

Intriguingly, several of the predicted AP-1/PGC-1α target genes are also under the control of PGC-1α working with other transcription factors. For example, the vascular endothelial growth factor (VEGF) or, based on the gene expression arrays, 8 OXPHOS genes seem likewise to be under the control of AP-1 as well as ERRα in the context of elevated PGC-1α in skeletal muscle (31, 36). We therefore assessed the predicted and experimental overlap of these two transcription factors in the regulation of AP-1/PGC-1α target genes. Interestingly, when the PCA analysis of the PGC-1α peaks was stratified in terms of eigenpeaks, we observed two distinct groups of peaks associated with AP-1/PGC-1α target genes (Fig. 5H). First, some of these genes exclusively harbored peaks with FOS-JUN-like TFBSs, whereas the second group exhibited either peaks with both FOS-JUN- and ESRRα-like TFBSs, or a
combination of distinct peaks with either of these sites within 10 kb from their promoters (Fig. 5H). Next, we validated this prediction by investigating the change in expression of different AP-1/PGC-1α target genes in the context of reduced ERRα expression and function, elicited by a combination of shRNA-mediated knockdown and pharmacological treatment of muscle cells with the ERRα inverse agonist XCT790 (31). In line with the PCA, two distinct groups of ERRα inhibition-sensitive (Fig. 5I-K) and –insensitive (Fig. 5L-N) AP-1/PGC-1α target genes were found.

Finally, since all of the experiments were performed in differentiated myotubes in culture, we assessed whether similar gene expression changes of the direct AP-1/PGC-1α targets involved in hypoxic gene regulation are also observed in skeletal muscle tissue of different gain- (6) and loss-of-function mouse models (7, 8) in vivo. In skeletal muscle-specific PGC-1α knockout mice, the expression of several of these genes was reduced significantly (Fig. 6A-F). Surprisingly however, some of the predicted transcripts were not altered in this loss-of-function model for PGC-1α, for example Nr0b2 (Fig. 6E). To further clarify the role of PGC-1α in the regulation of these genes, relative transcript levels were next assessed in muscle-specific transgenic mice for PGC-1α (Fig. 6G-L). In most cases, the genes with a reduction in their transcription in the PGC-1α muscle-specific knockout animals were inversely elevated in the PGC-1α muscle-specific transgenic mice. Moreover, some of these genes were likewise induced by exercise (Fig. 6G-L) and at least in some cases, for example Twf2 and Nr0b2 (Fig. 6J and K), PGC-1α overexpression and physical activity synergistically boosted gene expression, for Nr0b2 even in the absence of any effect of the muscle-specific PGC-1α transgene per se (Fig. 6K).
DISCUSSION

Exercise-induced skeletal muscle cell plasticity is a highly complex biological program that involves the remodeling of a number of fundamental cellular properties. Since PGC-1α function has been strongly linked to the induction of an endurance-trained muscle phenotype, we here dissected the PGC-1α-controlled transcriptional network in muscle cells. First, our results reveal a broad recruitment of PGC-1α to many different sites in the mouse genome (7512 peaks), the majority of which were either not located within 10 kb distance from a promoter or close to a gene that was not regulated by PGC-1α over-expression at the time of harvest of the cells, as has analogously been observed in many other ChIP-Seq experiments (for example, see ref. 37). Apart from the fact that PGC-1α could mediate long-range enhancer effects that were excluded in our peak-gene assignment, it is conceivable that PGC-1α recruitment is transcriptionally silent in some binding peaks because it requires the recruitment of additional cofactors for activation, which are not present in the conditions or cell type in which our experiments were performed. In addition, it is possible that a large fraction of PGC-1α binding peaks may be “neutral” in the sense of not having any direct role in regulating gene expression.

Second, while an almost equally strong effect of PGC-1α on gene induction and repression has been reported (31), our analysis now indicates that direct PGC-1α-mediated gene expression is restricted almost exclusively to positively regulated PGC-1α target genes, whereas the vast majority of gene repression is indirect, i.e. not associated with PGC-1α recruitment within a 10 kb distance to their promoters. Thus, the fact that almost 95% of all repressed genes were not linked to PGC-1α recruitment strongly implies that this coregulator primarily acts as a coactivator, and not as a corepressor as suggested by the data of some studies (38-40). Importantly, indirect repression of PGC-1α target genes was also supported by the MARA prediction. The strong indirect inhibition of genes, many of which are involved in inflammatory processes, is predicted by MARA to be mediated by TFs such as NFκB and IRF factors. Such an indirect inhibition of NFκB and pro-inflammatory genes by PGC-1α in muscle cells has been reported previously (33).

One of the main functions of PGC-1α in all cells and organs is to boost mitochondrial gene transcription and oxidative metabolism. Accordingly, we observed that Gene Ontology terms related to these pathways were highly enriched when analyzing positively regulated PGC-1α
target genes in muscle cells. Based on previous studies, the regulation of this core function could have been assigned to the direct interaction of PGC-1α and ERRα binding to regulatory elements of these genes (31, 32). Surprisingly, our data indicate that many of the genes that are involved in oxidative metabolic pathways are indirectly controlled by PGC-1α and, hence, do not require PGC-1α recruitment to enhancer and promoter elements. Likewise unexpectedly, the MARA analysis implies ERRα action on direct and indirect PGC-1α-induced target genes, i.e. in the presence or absence of PGC-1α coactivation. Thus, while these observations might obviously reflect a temporally distinct control of different PGC-1α target genes that is not represented in our simultaneous analysis of DNA binding and gene expression at one time point, it is conceivable that PGC-1α acts primarily as an upstream regulator of other factors that are subsequently controlling more downstream PGC-1α target genes without direct involvement of PGC-1α itself.

In skeletal muscle, PGC-1α has been reported to interact with ERRs, PPARs and other nuclear receptors, as well as myocyte enhancer and nuclear respiratory factors to mediate transcriptional regulation (3). Accordingly, ERRα and other nuclear receptor binding motifs were amongst the most highly significant binding elements in our present report. Importantly however, we also predict a number of so-far unknown TFs to functionally interact with PGC-1α and thereby contribute to PGC-1α-controlled gene expression in skeletal muscle. Since a complete functional validation of all new putative TF partners is beyond the scope of this manuscript, we combined the high-throughput results with several computational analyses (see Table 1) to select and test some of the potentially most important factors together with predicted target genes. Notably, in siRNA-based knockdown experiments, we could show that depletion of FOS and its putative AP-1 multimerization partners JUN and ATF3 as well as NFE2L2, ZFP143 and GTF2I in muscle cells reduced the ability of PGC-1α to positively regulate target genes. Second, we could provide evidence of a co-recruitment of FOS and PGC-1α to the same regulatory sites in the vicinity of AP-1/PGC-1α target genes, confirming a functional interaction between these TFs and PGC-1α. Thus, our results indicate that the coactivation repertoire of PGC-1α in muscle exceeds the prediction of previous studies by far. For example, even in our list of the top 15 motifs, several predicted TFs have not yet been investigated in the context of PGC-1α-controlled gene expression, including BPTF, FOSL2, REST or RREB1. Future studies will aim at a more
detailed dissection of the global functional consequences of PGC-1α coactivation of these TFs in muscle cells.

Curiously, almost all of our analyses, and in particular the principal component analysis, highlighted the relevance of FOS-JUN-like motifs. In fact, the largest amount of variation in TFBS occurrence within PGC-1α binding peaks results from either ESRRA-like or FOS-JUN-like motifs. The FOS-JUN-like motif, in particular, embodies the main binding elements of the AP-1 complex, which consists of different configurations of FOS, JUN, ATF and MAF proteins (35, 41). Our data comparing gene expression in cells with reduced FOS, JUN and ATF3 levels indicate that PGC-1α functionally interacts with the AP-1 complex in different configurations in the regulation of specific genes. The differential requirement observed for distinct AP-1 components might provide an additional layer of control for specific PGC-1α target gene regulation.

AP-1 function itself is regulated by a variety of stimuli, including cytokines, growth factors and stress, and subsequently controls a number of cellular processes including apoptosis, cell proliferation and differentiation, stress response and hypoxia (41, 42). Mechanistically, we classified PGC-1α-induced/AP-1-knocked-down targets in either direct or indirect genes. Most interestingly, functional analysis of these two groups of genes revealed that when AP-1 and PGC-1α act disjointedly, they are involved in the regulation of mitochondrial and other metabolic genes while, when coactivated by PGC-1α, AP-1 distinctly alters the expression of genes that are enriched in the ontology terms “response to hypoxia” and “mTOR signaling” (Fig. 5F). Intriguingly, a closer analysis of all 47 direct AP-1/PGC-1α target genes revealed 24 genes that are induced by hypoxia, are effectors of hypoxia or attenuate the detrimental consequences of hypoxia (Fig. 6M). For example, several inhibitors of the mTOR signaling pathways are included in this group of genes and hypoxia has been described as a suppressor of mTORC1 activity (43). Another group of genes contributes to the reduction of cellular stress, detrimental metabolites, reactive oxygen species and increase in cellular survival to reduce potential harmful consequences of prolonged hypoxia (44). Furthermore, several genes promote endothelial regeneration, vascular remodeling and vascularization (45). In this context, PGC-1α has previously been shown to promote VEGF-induced angiogenesis in skeletal muscle in a hypoxia-inducible factor 1α (HIF-1α)-independent, ERRα-dependent manner (36). Similarly, PGC-1α regulates the hypoxic response of brown fat (46),
neuronal and endothelial cells (47) even though the mechanisms of cellular protection exerted by PGC-1α in these experimental contexts have not been elucidated. Our findings now indicate that, to ensure adequate oxygen and nutrient supplies for oxidative metabolism in skeletal muscle cells, PGC-1α might coordinate metabolic needs through ERRα-induced Vegf expression with a broad, stress-induced AP-1-dependent hypoxia program. Such a functional convergence was found for a subset of the direct AP-1/PGC-1α target genes that likewise seem to be under the control of ERRα together with PGC-1α (Fig. 5H and I-K). Inversely, for the complementary subset of these genes, the functional interaction between AP-1 and PGC-1α seems distinct from the ERRα-dependent PGC-1α target gene regulation. Finally, in vivo evidence supports our muscle cell cultured-based prediction, considering that many of the AP-1/PGC-1α hypoxia-related target genes exhibit reduced and elevated transcript levels in PGC-1α muscle-specific knockout and transgenic animals, respectively. As previously demonstrated for VEGF and skeletal muscle vascularization (36), many aspects of the phenotypic consequences of exercise-induced muscle hypoxia occur in the muscle-specific transgenic mice even in the absence of physical activity. In extension of these studies, we now however found additional genes involved in this process that show an additional, or in case of Nr0b2, even an exclusive synergistic activation by exercise in the PGC-1α transgenic animals. Thus, combined with previous descriptions of muscle plasticity in these mice post-exercise in regard to insulin sensitivity (29), our present findings reiterate the importance of bona fide exercise even in a genetic model for endurance training such as the PGC-1α muscle-specific transgenic animals.

In summary, our data provide a first insight into the transcriptional network controlled by PGC-1α in muscle cells. While one other study of global DNA recruitment of PGC-1α has been performed in the human hepatoma cell line HepG2 (48), our results highlight the importance of combining ChIP-Seq experiment, transcriptional data together with a comprehensive computational modeling approach and experimental validation of predicted key regulators, in order to be able to discover mechanistic as well as functional outcomes of such a network. Combined with the knowledge of transcriptional regulation, posttranslational modifications, alternative splicing and recruitment of different chromatin remodeling protein complexes, a scenario can thus be conceived in which PGC-1α is able to control and integrate different signaling pathways using a multitude of different transcription factor binding partners (10, 11). A better understanding of such regulatory
networks will eventually allow the targeting of whole biological programs or specific sub-modules in pathological states of deregulation.
ACCESSION NUMBERS

The Gene Expression Omnibus (GEO) accession number for the ChIP-seq and gene expression array data reported in this paper is GSE51191.

ACKNOWLEDGMENTS

We would like to thank Dr. Anastasia Kralli, Svenia Schnyder, Gesa Santos, Kristoffer Svensson and Markus Beer for reagents, help and input for the preparation of this manuscript. This project was funded by the Swiss National Science Foundation (31003A_135397 to EvN, 310030_132900 to CH), SystemsX.ch (CellPlasticity, StoNets, and BrainstemX research projects to EvN), the Swiss Society for Research on Muscle Diseases (SSEM), the Neuromuscular Research Association Basel (NeRAB), the Gebert-Rüf Foundation “Rare Diseases” Program, the University of Basel and the Biozentrum. SS was supported by an IPPhD fellowship of the SystemsX.ch Swiss Initiative in Systems Biology. We are thankful of the [BC]2 Basel Computational Biology Center for providing computational resources.

Disclosure declaration: The authors have no conflict of interest in regard to this manuscript.
REFERENCES


43. **Cam H, Easton JB, High A, Houghton PJ.** 2010. mTORC1 signaling under hypoxic conditions is controlled by ATM-dependent phosphorylation of HIF-1alpha. Mol Cell 40:509-520.


FIGURE LEGENDS

**Figure 1. Genome-wide DNA recruitment of PGC-1α in mouse muscle cells.**

(A) PGC-1α ChIP-Seq binding peaks (read densities) around the TSS of the genes Acadm and Cycs obtained from the UCSC Genome Browser.

(B) Real-time PCR validation of the ChIP enrichment measured at the promoter of a set of PGC-1α target genes. Bars represent fold enrichment over that of the Tbp intron, error bars represent SEM. *p < 0.05; **p < 0.01; ***p < 0.001.

(C) Mapping ChIP-Seq PGC-1α peaks across the genome. Transcription Start Site (TSS) and Transcription End Site (TES) are relative to mm9 RefSeq transcripts. “Intergenic”: ≥ 10 kb from the nearest transcript; “Upstream of TSS”: -10 to 0 kb from the TSS; “Downstream of TES”: 0 to 10 kb from the TES. Numbers between brackets indicate, for each category, the ratio between the percentage of PGC-1α peaks and the percentage of the same number of randomly distributed peaks.

(D) Histogram illustrating the number of direct and indirect genes either up- or down-regulated by over-expression of PGC-1α in muscle cells. Direct genes are those associated to promoters found within ±10 kb relative to the nearest peak.

(E) Distribution of the distances of 532 peaks from their associated up-regulated gene promoters.

(F) Distribution of the distances of 43 peaks from their associated down-regulated gene promoters.

(G-H) Subset of the top significantly enriched GO Biological Process terms identified for directly and indirectly up-regulated (G) and down-regulated (H) PGC-1α target genes.

**Figure 2. Four distinct mechanistic modes of action for gene expression regulated by PGC-1α and TF partners.**

(A) Classification of direct and indirect target genes in MARA (see Methods)
(B) Distribution of peak distance from the closest promoter and phastCons conservation score of the peak.

(C) Distribution of log2 expression values for all mouse promoters. Expression values were averaged across the 3 GFP and the 3 PGC-1α samples. Direct targets are depicted in red, indirect targets in grey.

(D-G) Activity plot of the motifs ELF1,2,4 (D), ESRRA (E), REST (F) and NFKB1_REL_RELA (G) as predicted by MARA (Motif Activity Response Analysis). Red: direct targets; green: indirect targets.

Figure 3. PCA reveals FOS-JUN-like leucine zippers as a new class of putative functional PGC-1α partners.

(A) Sequence logo of the top position weight matrix discovered de novo by PhyloGibbs in the top 200 scoring peaks and of the corresponding canonical motif of ERRα as predicted by STAMP.

(B) Top scoring results of motif search performed on all 7512 PGC-1α peaks with MotEvo. Motifs depicted in red and blue correspond to the clusters identified by PCA in panel D.

(C) Top scoring results of motif search performed on the 3656 “non ESRRA-like” peaks with MotEvo.

(D) Fraction of explained variance of the top 10 PCA components.

(E) PCA analysis of the 7512 PGC-1α peaks. Eigenmotif scores across Principal Component 1 (PC1) and Principal Component 2 (PC2) are shown. Red and blue ellipses highlight motif clusters, as identified by PC1, of nuclear hormone receptor-like zinc finger and FOS-JUN-like leucine zipper proteins, respectively.

(F) Correlation between Principal Component 2 scores and binding site posterior sum for each peak relative to the top 10 PCA motifs. “r” refers to the Pearson correlation coefficient.
Figure 4. Validation of TFs associated with top scoring motifs reveals novel functional PGC-1α partners

(A) siRNA-mediated knockdown efficiency for FOS. Bars represent fold induction over GFP/siCtrl value, error bars represent SEM. *p < 0.05; **p < 0.01; ***p < 0.001. See also Figure S4.

(B-H) qRT-PCR analysis of PGC-1α target genes whose associated peak contains at least one binding site for the motif: FOS_FOS(B,L1)_JUN(B,D) (B-D), NFE2L2 (E), ZNF143 (F), GTF2I (G), NFY(A,B,C) (H). Bars represent % change compared to PGC-1α/siCtrl values. Error bars represent SEM. *p < 0.05; **p < 0.01; ***p < 0.001.

Figure 5. PGC-1α controls the hypoxia gene program via a functional interaction with different configurations of the AP-1 protein complex.

(A-C) qRT-PCR analysis of Cdk15, Nppb and Slc6a19 mRNA levels in response to PGC-1α overexpression and either siFos, siJun or siAtf3 knockdown. Data are normalized to mRNA levels in GFP infected cells. Error bars represent ± SEM. *p < 0.05; **p < 0.01; ***p < 0.001.

(D) Venn diagram illustrating the overlap in number of genes up-regulated by PGC-1α and down-regulated by either FOS, JUN or ATF3 knockdown.

(E) Histogram illustrating the number of direct and indirect PGC-1α/AP-1 target genes.

(F) Subset of the top significantly enriched Gene Ontology and KEGG terms identified for the two gene groups illustrated in panel (E).

(G) qRT-PCR validation of the ChIP enrichment of c-Fos measured at the gene of TGFβ1 (validated) and at the promoters of Nr0b2, Gprc5a and Dbt (predicted) target genes. Bars represent fold enrichment over PGC-1α exon2 set as 1. Error bars represent SEM. *p < 0.05; **p < 0.01; ***p < 0.001.

(H) PCA analysis of the 7512 PGC-1α peaks. Eigenpeak scores across Principal Component 1 (PC1) and Principal Component 2 (PC2) are shown. Colored dots correspond to peaks associated to the 47 direct PGC-1α/AP-1 targets. Blue dots refer to genes associated to
peaks containing only FOS-JUN TFBSs, while red dots refer to genes associated to peaks with FOS-JUN and ESRRA TFBSs, either located in the same peak or in distinct PGC-1α peaks.

(I-K) qRT-PCR analysis of PGC-1α/AP1 targets whose associated peaks contain an ESRRA binding site. The bars represent relative mRNA levels compared to AV-shGFP + AV-GFP + vehicle, which is set as 1. The error bars represent SEM. *p < 0.05, **p < 0.01, ***p < 0.001.

(L-N) qRT-PCR analysis of PGC-1α/AP1 targets whose associated peaks (if any) do not contain an ESRRA binding site. The bars represent relative mRNA levels compared to AV-shGFP + AV-GFP + vehicle, which is set as 1. The error bars represent SEM. *p < 0.05, **p < 0.01, ***p < 0.001.

Figure 6. PGC-1α controls the hypoxic gene program in muscle in vivo. (A-F) qRT-PCR analysis of hypoxic genes in sedentary control (ctrl) and muscle-specific knockout mice (MKO). The control group is set as 1. Error bars represent SEM. *p < 0.05, **p < 0.01, ***p < 0.001.

(G-L) qRT-PCR analysis of hypoxic genes in treadmill running mice. Control (ctrl) and muscle-specific transgenic (TG) mice were used under sedentary and exercise conditions. The control group sedentary is set as 1. Error bars represent SEM. *p < 0.05, **p < 0.01, ***p < 0.001. (M) Schematic representation depicting the downstream effects of the functional interaction between PGC-1α and the AP-1 complex in the context of the hypoxia gene program. Direct targets of PGC-1α and AP-1 are indicated in bold.
Table 1. Global summary of all analyses performed on PGC-1α peaks. The final score is the count of all analyses where a certain motif passed the defined cutoffs. The motifs chosen for validation and their corresponding values which satisfied the cutoffs are shown in bold.

<table>
<thead>
<tr>
<th>Motif name</th>
<th>PCA</th>
<th>Over-repr. in all PGC-1α peaks</th>
<th>Over-repr. in “non ESRRα-like” peaks</th>
<th>MARA activity Z score direct</th>
<th>MARA activity Z score indirect</th>
<th>Log2FC in expr. array</th>
<th>Abs. expr. in PGC-1α sample</th>
<th>Final ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESRRα</td>
<td>Yes</td>
<td>1</td>
<td>182</td>
<td>6.04(14.78)</td>
<td>15.49(37.94)</td>
<td>2.31</td>
<td>1829.45</td>
<td>6</td>
</tr>
<tr>
<td>FOS_FOS(B,L1)_JUN(B,D)</td>
<td>Yes</td>
<td>5</td>
<td>2</td>
<td>0.88(2.14)</td>
<td>1.81(-4.34)</td>
<td>1.78</td>
<td>1508.85</td>
<td>5</td>
</tr>
<tr>
<td>ZNF143</td>
<td></td>
<td>27</td>
<td>28</td>
<td>2.48(6.05)</td>
<td>4.65(9.68)</td>
<td>0.38</td>
<td>384.36</td>
<td>5</td>
</tr>
<tr>
<td>BPTF</td>
<td></td>
<td>21</td>
<td>12</td>
<td>1.38(3.37)</td>
<td>2.56(-6.25)</td>
<td>-0.56</td>
<td>333.34</td>
<td>4</td>
</tr>
<tr>
<td>ESRI</td>
<td></td>
<td>17</td>
<td>50</td>
<td>2.33(5.69)</td>
<td>4.53(11.04)</td>
<td>-0.47</td>
<td>232.42</td>
<td>4</td>
</tr>
<tr>
<td>FOSL2</td>
<td></td>
<td>6</td>
<td>3</td>
<td>0.88(2.14)</td>
<td>1.51(3.65)</td>
<td>-0.98</td>
<td>717.09</td>
<td>4</td>
</tr>
<tr>
<td>GTF2I</td>
<td></td>
<td>34</td>
<td>13</td>
<td>2.09(5.10)</td>
<td>2.38(-5.80)</td>
<td>-0.55</td>
<td>1207.81</td>
<td>4</td>
</tr>
<tr>
<td>NFE2L2</td>
<td>Yes</td>
<td>8</td>
<td>5</td>
<td>0.57(1.38)</td>
<td>1.01(-2.37)</td>
<td>-0.38</td>
<td>3673.63</td>
<td>4</td>
</tr>
<tr>
<td>NFY(A,B,C)</td>
<td></td>
<td>96</td>
<td>116</td>
<td>2.37(5.80)</td>
<td>3.56(7.62)</td>
<td>1.07</td>
<td>2409.48</td>
<td>4</td>
</tr>
<tr>
<td>NR5A1,2</td>
<td>Yes</td>
<td>3</td>
<td>188</td>
<td>3.53(8.66)</td>
<td>7.73(17.00)</td>
<td>-0.08</td>
<td>80.97</td>
<td>4</td>
</tr>
<tr>
<td>REST</td>
<td></td>
<td>12</td>
<td>6</td>
<td>0.48(1.15)</td>
<td>2.41(5.70)</td>
<td>-0.89</td>
<td>328.04</td>
<td>4</td>
</tr>
<tr>
<td>RREB1</td>
<td></td>
<td>15</td>
<td>10</td>
<td>1.56(3.82)</td>
<td>2.39(-5.42)</td>
<td>0.05</td>
<td>678.44</td>
<td>4</td>
</tr>
<tr>
<td>SP1</td>
<td></td>
<td>24</td>
<td>22</td>
<td>3.99(9.76)</td>
<td>0.61(0.33)</td>
<td>-0.32</td>
<td>751.98</td>
<td>4</td>
</tr>
<tr>
<td>STAT2,4,6</td>
<td></td>
<td>29</td>
<td>23</td>
<td>0.35(0.52)</td>
<td>4.81(-9.67)</td>
<td>-2.72</td>
<td>380.12</td>
<td>4</td>
</tr>
<tr>
<td>TLX1-3_NFIC(dimer)</td>
<td></td>
<td>19</td>
<td>17</td>
<td>0.84(-2.05)</td>
<td>4.91(-11.97)</td>
<td>-0.34</td>
<td>2339.33</td>
<td>4</td>
</tr>
</tbody>
</table>

a Requirement for PCA: being among the top 10 most contributing motifs to PC1 and PC2.
b Requirement for motifs over-representation: being among the top 30 significant motifs; ranking position shown.
c Requirement for MARA: have a Z-score ≥ 2.0. Numbers between brackets show the difference between the PGC-1α state and the GFP state, representing the direction in which the motif activity changes following PGC-1α over-expression.
d Requirement for the expression array (1): having a log2 fold change value ≥ 1.0 (corresponding to 2 folds up-regulation)
e Requirement for the expression array (2): having an absolute expression in the PGC-1α sample ≥ 100
**Fig. 1**

A. Distribution of PGC-1α peaks.

B. Fold enrichment over TBP intron.

C. Total number of peaks: 7512.

D. Number of genes: 341 direct, 643 indirect.

E. Distribution of peaks around the assigned promoter.

F. Density plot showing peak distribution.

G. Directly regulated genes:
- Gen. of prec. met. and energy (532 peaks).
- Oxidation-reduction pr.
- El. transport chain.
- Aerobic respiration.
- Tricarboxylic acid cycle.
- Acetyl-CoA catabolic pr.
- Acetyl-CoA metabolic pr.
- Cellular respiration.
- Organic acid metabolic pr.
- Carboxylic acid metabolic pr.

H. Indirectly regulated genes:
- Striated muscle contraction.

I. Differential regulation genes assignment.
- Number of genes: 0, 200, 600, 1000.
- Directly regulated genes:
  - (2 fold): 341.
  - (0.76 fold): 643.
  - (1.25 fold): 691.
  - (1.5 fold): 36.
- Indirectly regulated genes:
  - (1.24 fold): 341.

J. Gene ontology (GO) biological processes.
**Direct targets:**
- dist. < 1 kb
- dist. < 100 kb AND conservation > 50%

**Indirect targets:**
- all the other cases

![Diagram](image)

**Fig. 2**

---

**Panel A:**

**Panel B:**

**Panel C:**

**Panel D:**

**Panel E:**

**Panel F:**

**Panel G:**
A. Motif inferred by PhyloGibbs

B. Motif name | Z score | TFBSs count | Peaks percentage | Genes in the associated peaks
--- | --- | --- | --- | ---
ESRR | 96.95 | 1709 | 19.76% | 461 (28.47%)
BACH2 | 81.74 | 1300 | 16.01% | 286 (17.67%)
NRF1A,2 | 80.48 | 2007 | 22.58% | 522 (32.24%)
NRF2A | 72.63 | 707 | 8.24% | 200 (12.35%)
FOS_FOS(B,L1)_JUN(B,D) | 69.89 | 2030 | 19.18% | 341 (21.06%)
FOSL2 | 68.06 | 1360 | 16.05% | 308 (19.02%)
NFE2 | 62.78 | 1006 | 12.15% | 229 (14.14%)
NFE2L2 | 56.66 | 697 | 8.53% | 174 (10.75%)
HNF4A_NR2F1,2 | 47.20 | 766 | 9.15% | 207 (12.91%)
CTCF | 47.10 | 188 | 2.46% | 73 (4.51%)
PPARG | 40.61 | 389 | 4.30% | 101 (6.24%)
REST | 40.00 | 98 | 0.71% | 15 (0.93%)
NR1H4 | 38.91 | 596 | 7.08% | 156 (9.64%)
RXR_AVR(Dimer) | 35.02 | 364 | 4.26% | 100 (6.18%)
RREB1 | 34.23 | 180 | 1.78% | 42 (2.59%)

C. Motif name | Z score | TFBSs count | Peaks percentage
--- | --- | --- | ---
BACH2 | 59.05 | 712 | 16.96%
FOS_FOS(B,L1)_JUN(B,D) | 51.92 | 1125 | 20.18%
FOSL2 | 49.84 | 724 | 16.55%
NFE2 | 46.72 | 548 | 12.76%
NFE2L2 | 39.79 | 372 | 8.84%
REST | 35.42 | 71 | 0.93%
CTCF | 33.65 | 97 | 2.46%
CDX1,2,4 | 27.37 | 218 | 4.59%
MZF1 | 23.94 | 228 | 4.95%
RREB1 | 22.80 | 108 | 2.02%
SP1 | 21.90 | 216 | 5.50%
BPTF | 21.88 | 278 | 4.80%
GTF2I | 21.17 | 228 | 4.95%
PATZ1 | 20.97 | 162 | 3.92%
FOX1,2 | 20.47 | 211 | 4.33%

D. Fraction of Explained Variance

E. Principal Component 1 vs. Principal Component 2

F. PC2 score vs. TFBSs posterior sum
- Cellular stress response
- AMPK
- mTOR
- Cab39
- Ddit4
- Cell survival
- Oxidative stress reduction
- Vegfa
- Ciapin1
- Angiogenesis
- Vascularization
- Gabarapl1
- Autophagy
- Eef1a2
- protein quality control
- Arrb1
- Increased Vegfr activation
- Hagh
- Reduction of glycation products
- FA uptake
- Actin stress fiber remodelling
- Twf2
- PI3K
- Akt
- Inpp5j
- PGC-1α
- Acacb
- Fabp3
- Nppb
- Thiamine
- cell survival
- Reduced inflammation
- Sbno2
- Nr0b2
- Gprc5a
- Lonp1
- Efhd1
- Atp1b1
- Reduction of inflammation
- Ctrl sed
- Ctrl ex
- TG sed
- TG ex
<table>
<thead>
<tr>
<th>Motif name</th>
<th>Z direct</th>
<th>Z indirect</th>
<th>Z avg. direct</th>
<th>Z avg. indirect</th>
<th>Directly activated</th>
<th>Indirectly activated</th>
<th>Directly repressed</th>
<th>Indirectly repressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1: motifs only directly activated by PGC-1alpha</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP1.p2</td>
<td>3.99</td>
<td>0.61</td>
<td>9.76</td>
<td>0.33</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ELF1,2,4.p2</td>
<td>3.11</td>
<td>1.32</td>
<td>7.59</td>
<td>3.13</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PAX4.p2</td>
<td>2.50</td>
<td>1.53</td>
<td>6.11</td>
<td>-3.68</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LMO2.p2</td>
<td>2.36</td>
<td>1.65</td>
<td>5.78</td>
<td>3.98</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HNF4A_NR2F1,2.p2</td>
<td>2.26</td>
<td>1.54</td>
<td>5.52</td>
<td>3.64</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GTF2I.p2</td>
<td>2.09</td>
<td>2.38</td>
<td>5.10</td>
<td>-5.80</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Group 2: motifs directly and indirectly activated by PGC-1alpha</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESRRAR.p2</td>
<td>6.04</td>
<td>15.49</td>
<td>14.78</td>
<td>37.94</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NRSAP1.p2</td>
<td>3.53</td>
<td>7.73</td>
<td>8.66</td>
<td>17.00</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ZNF143.p2</td>
<td>2.48</td>
<td>4.65</td>
<td>6.05</td>
<td>9.68</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NPY[A,B,C].p2</td>
<td>2.37</td>
<td>3.56</td>
<td>5.80</td>
<td>7.62</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ESR1.p2</td>
<td>2.33</td>
<td>4.53</td>
<td>5.69</td>
<td>11.04</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RXR(A,B,G).p2</td>
<td>2.29</td>
<td>4.30</td>
<td>5.59</td>
<td>10.50</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Group 3: motifs only indirectly activated by PGC-1alpha</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NRF1.p2</td>
<td>1.60</td>
<td>4.61</td>
<td>3.91</td>
<td>6.21</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>YY1.p2</td>
<td>0.88</td>
<td>2.97</td>
<td>2.09</td>
<td>5.77</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>EHF;p2</td>
<td>0.73</td>
<td>2.77</td>
<td>1.77</td>
<td>6.35</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RXRA_VDR(dimer);p2</td>
<td>0.71</td>
<td>2.54</td>
<td>1.71</td>
<td>6.20</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HES1.p2</td>
<td>0.34</td>
<td>2.52</td>
<td>0.84</td>
<td>6.10</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>FOXO1,3,4.p2</td>
<td>0.46</td>
<td>2.51</td>
<td>1.13</td>
<td>6.12</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ELM1,4_GABP(A,B1);p3</td>
<td>1.18</td>
<td>2.46</td>
<td>2.89</td>
<td>5.95</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NFKB1REL_RELA.p2</td>
<td>0.50</td>
<td>6.54</td>
<td>1.19</td>
<td>-16.01</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Group 4: motifs only indirectly repressed by PGC-1alpha</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IRF1,2,7.p3</td>
<td>1.77</td>
<td>24.23</td>
<td>4.34</td>
<td>-14.48</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>MYOD1.p2</td>
<td>1.69</td>
<td>2.56</td>
<td>3.37</td>
<td>-6.25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>