OPEN

**Neuropsychopharmacology (2014) 39,** 2191–2199

**www.neuropsychopharmacology.org**

# A Molecular Profile of Cocaine Abuse Includes the Differential Expression of Genes that Regulate Transcription, Chromatin, and Dopamine Cell Phenotype

**Michael J Bannon*[,1], Magen M Johnson[1], Sharon K Michelhaugh[1], Zachary J Hartley[1], Steven D Halter[1], James A David[1], Gregory Kapatos[1,2] and Carl J Schmidt[3]**

*[1]Department of Pharmacology, Wayne State University School of Medicine, Detroit, MI, USA; [2]Center for Molecular Medicine & Genetics, Wayne State University School of Medicine, Detroit, MI, USA; [3]Department of Pathology, University of Michigan School of Medicine, Ann Arbor, MI, USA*

Chronic drug abuse, craving, and relapse are thought to be linked to long-lasting changes in neural gene expression arising through transcriptional and chromatin-related mechanisms. The key contributions of midbrain dopamine (DA)-synthesizing neurons throughout the addiction process provide a compelling rationale for determining the drug-induced molecular changes that occur in these cells. Yet our understanding of these processes remains rudimentary. The postmortem human brain constitutes a unique resource that can be exploited to gain insights into the pathophysiology of complex disorders such as drug addiction. In this study, we analyzed the profiles of midbrain gene expression in chronic cocaine abusers and well-matched drug-free control subjects using microarray and quantitative PCR. A small number of genes exhibited robust differential expression; many of these are involved in the regulation of transcription, chromatin, or DA cell phenotype. Transcript abundances for approximately half of these differentially expressed genes were diagnostic for assigning subjects to the cocaine-abusing *vs* control cohort. Identification of a molecular signature associated with pathophysiological changes occurring in cocaine abusers' midbrains should contribute to the development of biomarkers and novel therapeutic targets for drug addiction.

*Neuropsychopharmacology* (2014) **39,** 2191–2199; doi:10.1038/npp.2014.70; published online 16 April 2014

## INTRODUCTION

Drug addiction exacts a staggering toll on affected individuals and society as a whole. The persistence of drug craving and risk of relapse so characteristic of addiction are thought to be linked to long-lasting changes in neural gene expression arising through changes in transcription and chromatin regulation. In essence, such mechanisms are believed to constitute a 'molecular memory' that contributes to the maintenance of the drug-addicted state (Feng and Nestler, 2013).

Of the many different neural cell types and circuits implicated in the effects of drugs of abuse, perhaps none have a more central role than dopamine (DA)-synthesizing neurons of the ventral midbrain, which innervate widespread regions of the forebrain. Midbrain DA cells constitute a mere 1 in 200,000 neurons of the human brain, yet DA is critically involved in mediating both the acute rewarding effects of drugs of abuse and the conditioned responses to cues associated with previous drug use (Volkow *et al*, 2011). Conversely, chronic drug abuse may lead to deficits in DA neurotransmission that contribute to the emergence of adverse consequences (eg, anhedonia and dysphoria) upon cessation of drug use (Koob and Volkow, 2010). The critical contributions of DA neurons throughout the addiction process provide a compelling rationale for elucidating drug-induced molecular changes in these cells, yet our understanding of the nature of such changes (in either humans or model systems) remains far from complete.

Despite the challenges associated with its use, postmortem human brain constitutes a unique resource that can be exploited to develop new insights into the pathophysiology of complex disorders such as drug addiction (McCullumsmith *et al*, 2013). Data from human postmortem studies can help us to understand the extent to which animal models recapitulate key features of human disorders and, at the same time, lead to new hypotheses testable in model systems. Given the dearth of knowledge regarding the molecular effects of cocaine on DA neurons, and evidence of some important species differences in DA cell biology and susceptibility to age-related and drug-induced effects (Bannon and Whitty, 1997; Björklund and Dunnett 2007; Whitty *et al*, 1997), we conducted a discovery-driven analysis of the profile of gene expression in the human midbrain associated with chronic cocaine

*Correspondence: Dr MJ Bannon, Department of Pharmacology, Wayne State University School of Medicine, 6374 Scott Hall, 540 E. Canfield, Detroit, MI 48201, USA, Tel: +1 313 993 4271, Fax: +1 313 577 6739, E-mail: mbannon@med.wayne.edu
Received 11 December 2013; revised 19 February 2014; accepted 9 March 2014; accepted article preview online 19 March 2014

abuse. A molecular signature diagnostic of cocaine abuse was identified, which included differential expression of genes regulating transcription, chromatin, and DA cell phenotype. A better understanding of the core pathophysiological changes associated with cocaine abuse should lead to the development of biomarkers and novel therapeutic targets for drug addiction.

## MATERIALS AND METHODS

### Subject and Specimen Characterization

De-identified postmortem human brain specimens were collected during the routine autopsy process as described in detail previously (Albertson *et al*, 2004; Albertson *et al*, 2006; Bannon and Whitty, 1997; Johnson *et al*, 2011, Johnson *et al*, 2012; Zhou *et al*, 2014). In brief, the cause and manner of death were determined by forensic pathologists following medico-legal investigations evaluating the circumstances of death including medical records, police reports, autopsy results, and toxicological data (Karch, 2002). Subject inclusion in the cocaine cohort ($n=10$) was based on determination of cocaine abuse as the cause of death, a documented history of drug abuse, and a positive toxicology for cocaine and/or the cocaine metabolite benzoylecgonine but negative for other drugs of abuse or CNS medications at time of death. No significant correlation was observed between levels of cocaine or its metabolites and the abundance of individual transcripts studied (not shown). Control subjects ($n=10$) died as a result of cardiovascular disease or gunshot wound, had no documented history of drug abuse, and tested negative for cocaine and other drugs of abuse (with the exception of a single subject with a subintoxicating level (0.06 g/dl) of ethanol). Samples were not screened for the presence of nicotine or metabolites. Exclusion criteria for the study included a known history of neurological or psychiatric illness, death by suicide, estimated postmortem interval exceeding 20 h, evidence of neuropathology (eg, encephalitis or stroke), or chronic illness (eg, cirrhosis, cancer, HIV, or prolonged hospitalization). To reduce the variance between groups unrelated to drug abuse, each cocaine abuse subject was matched to a control subject based on race, gender, and age before inclusion in the study (most study subjects were African American and all were males). The final groups (Table 1) did not differ with regard to any of these parameters or with regard to well-established measures of sample quality and perimortem agonal state (ie, brain pH and RNA integrity number (RIN); Schroeder *et al*, 2006; Stan *et al*, 2006).

The methodologies used for the analyses of gene expression have also been described previously (Johnson *et al*, 2011; Johnson *et al*, 2012). In brief, intact brains were sectioned transversely at the level of the posterior diencephalon and mid-pons to obtain a tissue block encompassing the entire human midbrain (corresponding approximately to plates 51–56 of DeArmond *et al*, 1989). Fresh-frozen blocks were subsequently cryostat-sectioned to a thickness of 250 μm and slide-mounted. Guided by the presence of neuromelanin, DA cell-enriched regions of the substantia nigra (including the dorsal and ventral tiers, pars lateralis) and ventral tegmental area (including the interfascicular, central linear, paranigral, and parabrachial nuclei), identified as previously described (Bannon and Whitty, 1997; Whitty *et al*, 1997), were finely dissected and pooled into a single sample from each subject. RNA was isolated, quantified, and assessed for integrity. A pilot microarray experiment comparing this dissection method with samples obtained using a more standard block dissection of ventral midbrain found an enrichment of DA cell-specific transcripts (eg, TH: 2.7-fold) but not of transcripts indicative of glutamate neurons (VGLUT2: 0.2-fold), GABA neurons (VGAT: 1.2-fold), neurons overall (NSE: 0.7-fold), astroglia (GFAP: 1.2-fold), oligodendrocytes (MBP: 1.0-fold), or microglia (CD68: 1.0-fold).

Microarray procedures were performed at the Keck microarray facility (Yale Center for Genome Analysis) as previously described (Johnson *et al*, 2011; 2012). HT-12 BeadChips (Illumina, San Diego, CA) were hybridized with cRNAs generated from each specimen, scanned on an Illumina Iscan, and loaded into Illumina BeadStudio to evaluate spiked-in controls before quantile normalization of the data. Raw and normalized data have been deposited in the NCBI-GEO repository (GSE54839). Normalized microarray data were imported into MultiExperiment Viewer (MeV) (http://www.tm4.org/mev/). Hierarchical clustering of the expression profiles of sample replicates intentionally processed at different times ruled out microarray chip and batch effects (Supplementary Figure 1). Subsequent hierarchical clustering of all subject samples (20 samples $\times$ technical triplicates $=60$ total profiles) revealed only three potential outliers (one triplicate each from three cases); after their removal, the remaining replicates (17 triplicates and 3 duplicates) clustered precisely according to the sample of origin (Pearson $r=0.99$, two-tailed $P<0.0001$) and were thus averaged to obtain a single expression profile for each subject. A total of 16,301 of the 48,761 probes represented on the microarray were detected ($P\leqslant 0.05$) in the majority of subjects analyzed ($\geqslant 11/20$). A subset of differentially expressed genes identified by microarray was subsequently validated by quantitative real-time polymerase chain reaction PCR (qPCR) as previously described (Johnson *et al*, 2011; 2012)(primer sequences listed in Supplementary Table 1).

### Statistical and Bioinformatic Analyses

Group-wise differences in gene expression were assessed in MeV using a nonparametric two-class unpaired Rank Product analysis with a false discovery rate (FDR) multiple testing correction ($P\leqslant 0.05$, 5% FDR; Breitling *et al*, 2004). On the basis of our previous experience (Johnson *et al*, 2011; 2012), a 1.4-fold change cutoff was then applied to identify the most robust cocaine-related differential gene expression with the highest likelihood of validation. Annotations for differentially expressed probes defined in this manner were confirmed through BLAT sequence alignments in the UCSC Genome Browser (Human GRCh37/hg 19 assembly; http://genome.ucsc.edu/cgi-bin/hgGateway); probes with $\geqslant 48/50$ bp aligning to a single locus were further investigated. Microarray data for the resulting list of genes with robust differential expression (all of which were detected in 20/20 subjects) were imported into SPSS (IBM, Armonk, NY) for receiver operating characteristic (ROC)

**Table 1** Characteristics of Study Subjects

| | Control subjects | | | | | | | Cocaine abuse subjects | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Pair** | **Age** | **Race/Sex** | **COD** | **pH** | **RIN** | **Coc (μg/ml)** | **BE (μg/ml)** | **Age** | **Race/Sex** | **COD** | **pH** | **RIN** | **Coc (μg/ml)** | **BE (μg/ml)** |
| 1 | 47 | WM | Acute MI, ASCVD | 6.2 | 6.4 | ND | ND | 46 | WM | Acute cocaine intoxication | 6.4 | 7.3 | 0.8 | 5.7 |
| 2 | 45 | BM | MGSW | 6.3 | 6.1 | ND | ND | 46 | BM | Cocaine abuse | 6.6 | 6.1 | 0.069 | 1.5 |
| 3 | 49 | BM | MI, ASCVD | 6.8 | 7.5 | ND | ND | 49 | BM | Aortic aneurysm due to cocaine abuse | 6.4 | 6.9 | ND | 0.42 |
| 4 | 53 | BM | ASCVD | 6.3 | 7 | ND | ND | 59 | BM | Cocaine abuse | 6.6 | 6.5 | ND | 0.07 |
| 5 | 51 | BM | Aortic dissection, hypertension | 6.7 | 5.7 | ND | ND | 54 | BM | Cocaine abuse | 6.6 | 4.8 | ND | 0.09 |
| 6 | 46 | BM | MGSW | 6.8 | 7.5 | ND | ND | 52 | BM | Cocaine abuse | 6.6 | 7.6 | 0.078 | 0.33 |
| 7 | 45 | WM | GSW | 6.5 | 5.6 | ND | ND | 46 | WM | Cocaine abuse | 6.4 | 4.4 | ND | 0.03 |
| 8 | 52 | BM | Hypertensive cardiomyopathy | 6.3 | 7 | ND | ND | 52 | BM | Aortic dissection due to cocaine abuse | 6.3 | 7.2 | ND | 0.58 |
| 9 | 50 | BM | GSW | 6.7 | 6.9 | ND | ND | 52 | BM | Cocaine abuse | 6.5 | 6.7 | 0.29 | 3 |
| 10 | 45 | BM | ASCVD | 6.4 | 6.5 | ND | ND | 45 | BM | Cocaine abuse | 6.9 | 6.1 | 0.053 | 3.4 |
| $\bar{x}$ (SEM) | 49 (1) | | | 6.5 (0.1) | 6.6 (0.2) | | | 50 (1) | | | 6.5 (0.1) | 6.4 (0.3) | | |

Abbreviations: ASCVD, arteriosclerotic cardiovascular disease; BE, the major cocaine metabolite benzoylecgonine; BM, black male; Coc, cocaine; COD, cause of death; GSW, gunshot wound; MGSW, multiple gunshot wounds; MI, myocardial infarction; ND, not detected; RIN, RNA integrity number; SEM, standard error of the mean; WM, white male; $\bar{x}$, mean.
N.B.: subjects were added to the study in a matched pairwise manner but analyzed groupwise.

curve analysis (Metz, 2006), in which the true-positive rate (sensitivity) is plotted against the false-positive rate (100-specificity) for different cutoff points of transcript abundance. Calculated *P*-value reflects the likelihood that the area under an ROC curve is significantly different from chance in distinguishing between diagnostic groups (ie, cocaine users *vs* controls).

David (http://david.abcc.ncifcrf.gov/home.jsp) was used to identify Gene Ontology (GO) biological processes (GO_TERM_FAT) over-represented (operationally defined as containing $\geqslant 3$ gene list members and significant at $P \leqslant 0.01$ by the modified Fisher Exact test) in the list of cocaine-responsive genes in comparison with all human genes. In a similar manner, enrichment of transcription factor-binding sites in cocaine-responsive genes in comparison with all known genes was determined using the publically available cREMaG program (http://www.cremag.org/) (analyzing sequences 10 kb upstream to 5 kb downstream of transcriptional start sites, with an imposed cutoff of $P \leqslant 0.01$).

### Immunohistochemistry and Microscopy

Immunohistochemistry was carried out on thaw-mounted post-fixed tissue sections (14 μM) of the human ventral midbrain as previously described (Bannon *et al*, 2004), using overnight incubations with primary antisera directed against DAT (Millipore AAB1591P 1:500), FOXA2 (Millipore 07-633 1:500), TH (Dr. Greg Kapatos 1:50,000), c-FOS (Santa Cruz SC-52 1:100), MRAP2 (Novus NBP1-88752 1:200), CDKN1A (Santa Cruz SC-397 1:100), CCL2 (Abgent AP6699b 1:500), GADD45b (Santa Cruz CS-33172 1:200), SERPINA3 (Sigma HPA002560 1:100), or EGR1 (Santa Cruz SC-101033 1:500), followed by brief incubations with a biotinylated secondary antibody (Vector 1:500) and protein visualization using a diaminobenzidine peroxidase substrate kit with nickel enhancement (Vector). Signal specificity was confirmed in each case by the omission of primary antisera. Images were captured using an Olympus BX53 microscope under $\times 40$ objective and CellSens digital software with deconvolution.

## RESULTS

Postmortem specimens of the DA cell-enriched ventral midbrain from chronic cocaine abusers and well-matched, drug-free controls (Table 1) were processed in parallel, and all array-related procedures (eg, RT, cRNA generation, and array hybridization) performed in triplicate for each subject to obtain accurate estimates of gene expression profiles (see more details, Materials and Methods). Application of a statistical threshold ($P \leqslant 0.05$, FDR = 5%) and a magnitude of difference threshold ($\geqslant$1.4-fold difference) to profiles of gene expression resulted in a list of 91 genes (98 array probes) that exhibited robust, statistically significant differential expression between chronic cocaine abusers and matched control subjects (Table 2). Differential gene expression was validated by qPCR in every instance examined ($n = 11$), independent of transcript abundance (3- to 95-fold above background) or the magnitude of differences seen (3.6-fold increases to 2-fold decreases) (Figure 1), supporting the validity of the larger dataset of genes found to be differentially expressed in cocaine abusers.

To further characterize this dataset, a number of genes were selected for immunohistochemical localization of the encoded proteins. Proteins previously associated with the DA cell phenotype were found to be robustly and selectively expressed in midbrain DA neurons (also readily identifiable by their characteristic large nuclei and high intracellular neuromelanin content): tyrosine hydroxylase (TH) and the

**Table 2** Transcripts Differentially Expressed in the Ventral Midbrain of Cocaine Abusers

| Transcripts up-regulated 3.6 to 1.6 fold | | | | Transcripts up-regulated 1.5 to 1.4 fold | | | | Transcripts down-regulated −1.4 to −2.0 fold | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ILMN PROBE_ID | Gene symbol | Cocaine fold-difference | ROC sig. | ILMN PROBE_ID | Gene symbol | Cocaine fold-difference | ROC sig. | ILMN PROBE_ID | Gene symbol | Cocaine fold-difference | ROC Sig. |
| 1762899 | EGR1 | 3.6 | | 2129161 | LRRC32 | 1.5 | | 2325337 | APOL2 | −1.4 | √ |
| 1669523 | FOS | 3.3 | | 1719695 | NFKBIZ | 1.5 | √ | 1666385 | CALM3 | −1.4 | √ |
| 1784602 | CDKN1A | 3.2 | | 1773154 | NFKBIA | 1.5 | | 1706301 | RET | −1.4 | √ |
| 1720048 | CCL2 | 3.0 | √ | 1719599 | SYTL4 | 1.5 | | 1665107 | ITGB1BP1 | −1.4 | |
| 2374865 | ATF3 | 2.6 | | 2336781 | SOD2 | 1.5 | | 1712430 | ATP5G1 | −1.4 | √ |
| 1789074 | HSPA1A | 2.6 | | 1711566 | TIMP1 | 1.5 | | 1720482 | CEND1 | −1.4 | √ |
| 1801616 | EMP1 | 2.5 | √ | 1677402 | C11orf96 | 1.5 | | 1668052 | FOXA2 | −1.4 | √ |
| 1770338 | TM4SF1 | 2.5 | | 1702231 | C1orf54 | 1.5 | | 1801302 | SCN1B | −1.4 | |
| 1660436 | HSPA1B | 2.4 | | 1766713 | HSPD1 | 1.5 | | 1712461 | CBLN1 | −1.4 | |
| 1659766 | BAG3 | 2.3 | | 1803988 | MCL1 | 1.5 | √ | 1798177 | CHURC1 | −1.4 | |
| 1775304 | DNAJB1 | 2.3 | | 1778924 | PDE1A | 1.5 | | 1870705 | RAB3C | −1.4 | |
| 1806023 | JUN | 2.3 | √ | 1695706 | H3F3B | 1.5 | √ | 1677273 | TH | −1.4 | |
| 1751607 | FOSB | 2.2 | | 1774077 | GBP2 | 1.5 | √ | 1669788 | NUDT14 | −1.4 | √ |
| 1718977 | GADD45B | 2.2 | | 2053415 | LDLR | 1.5 | | 2356578 | TH | −1.4 | √ |
| 1767556 | DEPP | 2.1 | | 1781952 | MGST1 | 1.5 | | 2041236 | RAP1GAP2 | −1.4 | √ |
| 1703123 | CSRNP1 | 2.0 | | 1732410 | SLC16A9 | 1.4 | | 1684694 | ANK1 | −1.5 | √ |
| 1729188 | HAMP | 2.0 | | 1687508 | ALDH7A1 | 1.4 | | 12049274 | ARHGDIG | −1.5 | √ |
| 1782050 | CEBPD | 1.9 | | 1707339 | BTG3 | 1.4 | √ | 1794829 | MRAP2 | −1.5 | √ |
| 2132982 | IGFBP5 | 1.9 | | 1753342 | SAT1 | 1.4 | √ | 1808226 | RGS16 | −1.5 | √ |
| 1659936 | PPP1R15A | 1.8 | | 1804735 | CBS | 1.4 | | 1805376 | KCNJ6 | −1.5 | √ |
| 1741021 | CH25H | 1.8 | √ | 1658356 | PAMR1 | 1.4 | | 1701483 | SYP | −1.5 | √ |
| 1674236 | HSPB1 | 1.7 | | 1791890 | SPON1 | 1.4 | | 2105797 | SLC6A3 | −1.5 | √ |
| 1720829 | ZFP36 | 1.7 | | 2355168 | MGST1 | 1.4 | | 1658679 | YJEFN3 | −1.5 | √ |
| 1760347 | SRGN | 1.7 | | 2347068 | MKNK2 | 1.4 | | 2325338 | APOL2 | −1.5 | √ |
| 1704730 | CD93 | 1.7 | √ | 1667825 | MLKL | 1.4 | √ | 2397842 | SNCB | −1.5 | |
| 1788874 | SERPINA3 | 1.7 | √ | 1812666 | DNAJC15 | 1.4 | √ | 2333219 | ACHE | −1.5 | √ |
| 2169152 | SRGN | 1.7 | | 1768973 | HIST2H2AC | 1.4 | √ | 1815308 | SDC1 | −1.6 | √ |
| 1782788 | CSDA | 1.7 | | 1661599 | DDIT4 | 1.4 | | 1779343 | SNCB | −1.6 | √ |
| 1762255 | GSTM1 | 1.7 | | 1679268 | PELI1 | 1.4 | √ | 1652540 | RELL2 | −1.6 | √ |
| 2307903 | VCAM1 | 1.7 | √ | 1675448 | ZFP36L1 | 1.4 | | 2397846 | SNCB | −1.6 | |
| 2144426 | HIST2H2AA3 | 1.6 | √ | 1791576 | CHSY1 | 1.4 | √ | 1791726 | TUBB3 | −1.6 | √ |
| 1723522 | APOLD1 | 1.6 | | | | | | 2069224 | PVALB | −2.0 | √ |
| 2233539 | SLC39A8 | 1.6 | √ | | | | | | | | |
| 1750324 | IGFBP5 | 1.6 | | | | | | | | | |
| 1653028 | COL4A1 | 1.6 | √ | | | | | | | | |

Abbreviations: ROC, receiver operating characteristic.
Gene transcripts ordered by magnitude of differential expression between cocaine abusers and control subjects as determined by microarray analysis (separated into multiple columns for ease of visualization). ROC-significant ($p \leqslant 0.05$) transcripts predictive of subject assignment to cocaine cohort are indicated.

DA transporter (DAT), for example, were found throughout the soma and processes of DA cells (Figure 2a and b), whereas the transcription factor forkhead box A2 (FOXA2) showed a primarily nuclear localization (Figure 2c). A number of other differentially expressed genes for which there was no *a priori* information regarding locus of expression showed enriched expression within DA neurons to varying degrees (Figures 2d–j); none of the examined showed selective labeling of non-DA cells or glia (not shown).

In order to gauge whether differentially expressed genes might serve as potential biomarkers of cocaine abuse, we evaluated our dataset using receiver operating characteristic (ROC) curve analysis, a method commonly used for assessing the diagnostic performance (ie, the sensitivity and specificity) of assays (Metz, 2006). As shown in Table 2, numerous upregulated genes were found to be ROC significant ($P \leqslant 0.05$). Downregulated genes, though generally modest in terms of magnitude of difference, showed an even higher proportion of ROC significance (Table 2). Overall, approximately half of all the 91 genes with robust differential expression performed significantly better than chance ($P \leqslant 0.05$ to $P \leqslant 0.001$; 76–93% accuracy) in correctly assigning subjects to the cocaine abuse or the control cohort.

It is possible that a coordinate regulation of cocaine-responsive genes could arise through the actions of a few specific transcription factors acting on cognate binding

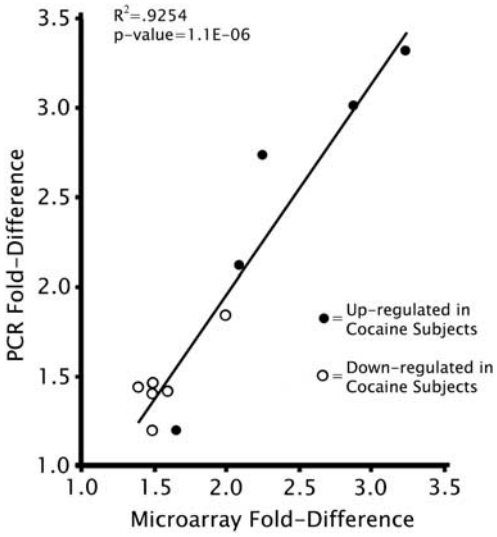


**Figure 1** Validation by qPCR of differences in transcript abundance identified by microarray. Note the significant correlation between the two measures, irrespective of the magnitude or direction of differential expression. Upregulated genes (largest to smallest increase): CDKN1A, CCL2, JUN, GADD45B, GSTM1. Downregulated genes (largest to smallest decrease): PVALB, SLC6A3, KCNJ6, MRAP2, TH, FOXA2. For primer information, see Supplementary Table 1.

sites in target genes. Using bioinformatics, we identified binding sites for several transcription factors that were significantly over-represented in cocaine-responsive genes compared with all human genes. The enrichment of GATA-binding sites (Figure 3a) is noteworthy because, although GATA expression was unchanged, expression of the GATA-binding partner and functional modulator CEBPD (Tong *et al*, 2005) was nearly doubled in cocaine abusers (Table 2). The over-representation of AP-1 sites (Figure 3a) is consonant with the striking induction of the immediate early genes (IEGs) FOS, FOSB, and JUN (Table 2), as these factors activate transcription through heterodimerization at AP-1 sites. Although STAT3 expression was unchanged in cocaine abusers, the significant enrichment of STAT-binding sites (Figure 3a) is consistent with changes in JAK2–STAT signaling reported in the DA neurons of rodents chronically exposed to cocaine (Berhow *et al*,1996). Furthermore, STAT3 elements mediate some effects of NFKB (Dhar *et al*, 2013), a transcription factor whose functional activity is most likely perturbed by the upregulation in NFKB modulators NFKBI and NFKBIZ (Table 2). Overall, the significant over-representation of GATA, AP-1, and STAT3 DNA elements in cocaine-responsive genes suggests roles for the associated transcription factors in coordinating cellular responses to cocaine.

Cocaine-responsive genes were also examined for evidence of their involvement in shared biological processes. Hierarchical clustering of subjects' gene expression profiles revealed that cocaine-responsive genes clustered into a number of distinct subgroups. The GO-defined biological processes significantly enriched in these discrete clusters are summarized in Figure 3b. Upregulated genes fall into three large clusters: Cluster A, characterized by chromatin organization, included several histone genes. Cluster B, which encompassed GO terms for behavior, drug response,

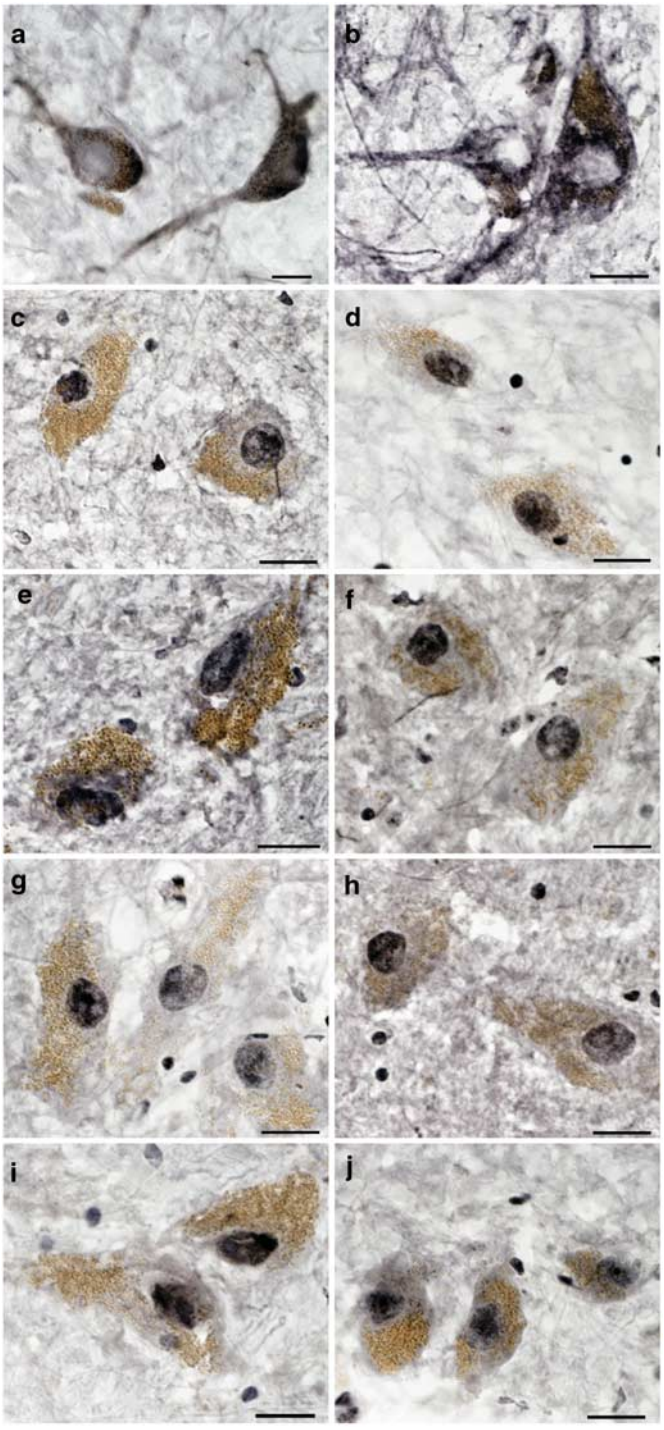


**Figure 2** Immunohistochemical localization of encoded proteins for several differentially expressed genes. Protein immunoreactivity was visualized as a dark purple precipitate. (a) TH, (b) DAT, (c) FOXA2, (d) MRAP2, (e) CCL2, (f) SERPINA3, (g) CDKN1A, (h) GADD45B, (i) FOS, and (j) EGR1. Note the prominent labeling of DA neurons, also readily identifiable by their characteristic large nuclei and intracellular neuromelanin accumulation (brown). Scale bars = 20 μm.

and transcriptional regulation, included a number of well-known cocaine-responsive transcription factors (eg, IEGs). Closely related clusters C, D, and E, which encompassed negative regulation of cell death, unfolded protein response,

**a**

| TFBS | Fold Over-Representation | p-value |
|---|---|---|
| GATA | 6.7 | 0.0009 |
| AP1 | 6.4 | 0.0023 |
| STAT3 | 6.1 | 0.0038 |

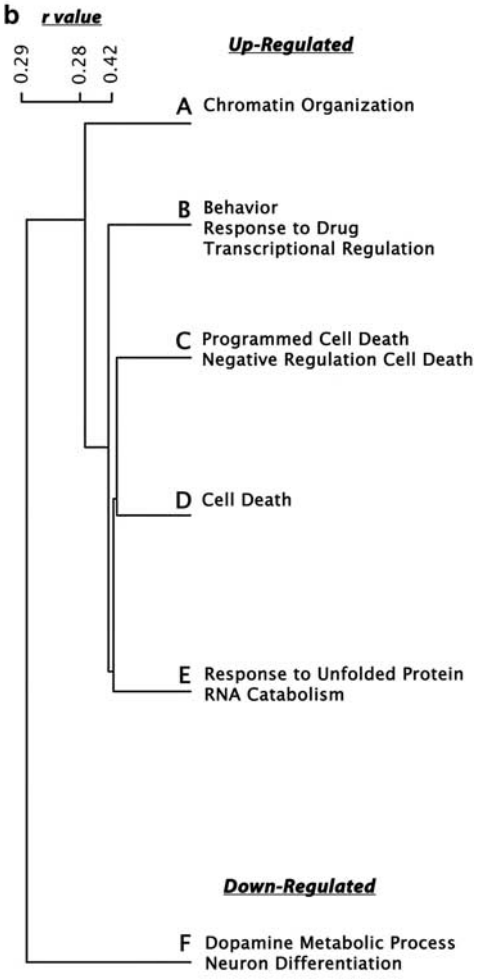


**Figure 3** Transcription factor binding sites and biological processes enriched in cocaine-responsive genes. (a) Binding sites for the transcription factors GATA, AP-1, and STAT3 were significantly enriched in genes with robust differential expression in cocaine abusers (Table 2) compared with all human genes (fold-enrichment and *P*-values indicated). (b) Cocaine-responsive genes (Table 2) hierarchically cluster into a small number of broad groups significantly enriched for distinct biological processes. Clusters A, B and C-D-E encompass upregulated genes, while cluster F encompasses all downregulated genes; *r* indicates correlation coefficients. More detailed hierarchical clustering and GO term enrichment data are provided in Supplementary Figure 2 and Supplementary Table 2, respectively. TFBS, transcription factor binding site.

and RNA catabolism, included heat-shock proteins and numerous other stress-response genes. In contrast, all downregulated genes fell within a single cluster (F), characterized by DA metabolic processes and neuronal differentiation, and which included numerous DA cell phenotypic genes. (Detailed hierarchical clustering and GO term enrichment data are provided in Supplementary Figure 2 and Supplementary Table 2, respectively.)

## DISCUSSION

The major goal of this study was to identify the most robust differences in midbrain gene expression associated with chronic cocaine abuse. Of the 91 genes exhibiting such differential expression (Table 2), upregulated genes fell into several distinct biological processes; prominent among these were chromatin and transcriptional regulation. Chromatin regulation might arise through the observed differential expression of histone-encoding genes HIST2H2AA3, HIST2H2AC and H3F3B. The latter is of particular interest, as it encodes a replication-independent H3 variant expressed primarily in non-dividing cells and is thought to be deposited at sites of nucleosomal displacement throughout transcribed genes, thus representing an epigenetic imprint of transcriptionally active chromatin (McKittrick *et al*, 2004). Cocaine abuse also had an impact on the expression of growth arrest and DNA damage-inducible beta (GADD45B) and the cyclin-dependent kinase inhibitor 1A (CDKN1A), genes originally described in the context of cell stress and cell cycle arrest. More recently, GADD45B has been a focus of intensive interest in neuroscience as an activity-dependent mediator of neurogenesis and synaptic activity acting via DNA demethylation of neural gene promoters, with CDKN1A as an abundant neural GADD45B-binding partner (Ma *et al*, 2009; Ploski *et al*, 2006). In the present study, GADD45B and CDKN1A were robustly expressed in human DA neurons under basal conditions (Figure 2) and substantially upregulated in cocaine abusers (Table 2; Figure 1). Increased capacity for DNA demethylation of targeted gene promoters in these subjects could contribute to the activation of other, downstream cocaine-responsive genes.

A large number of upregulated genes encode transcription factors or associated proteins (Table 2). Most previous studies of the regulation of transcription factors in drug abuse have focused on the forebrain, although inductions of c-fos and egr1 have also been observed in the midbrain of animals exposed to cocaine or cocaine-associated cues (Kufahl *et al*, 2009; Thomas *et al*, 2003). It is noteworthy that rodent studies generally report low basal levels of IEG expression and a diminution of induction with progression from acute to repeated drug exposure (Moratalla *et al*, 1996); in contrast, we found not only a substantial basal expression of numerous IEGs in human DA neurons (Figure 2) but also a robust upregulation of expression associated with recent cocaine use (as evidenced by positive toxicology) by chronic cocaine abusers (Table 2). In a similar vein, the persistent induction of IEGs and other transcription factors has been demonstrated in human neocortical foci prone to recurrent seizures (Beaumont *et al*, 2012), suggesting some unique facets of IEG regulation in the human brain that warrant further study. As mentioned above, the strong upregulation of numerous transcription factors (eg, FOS, FOSB, JUN, ATF3, CEPBD, NFKBI, and NFKBIZ) probably contributes to the cocaine-responsiveness of other, downstream genes that possess associated DNA response elements. In addition to these

well-characterized transcription factors, a number of other genes upregulated in our cocaine cohort (eg, CSRNP1, ZFP36, CSDA, GBP2; Table 2) encode proteins that modulate gene expression through a variety of transcriptional and/or post-transcriptional mechanisms (Franks and Lykke-Andersen 2007; Gingras *et al*, 2007; Hurt *et al*, 2004). Overall, many of the most robustly upregulated genes are positioned to mediate broader downstream effects on gene expression through changes in histone isoform or abundance, DNA demethylation, gene transcription, or RNA stability or export.

Very few of these upregulated genes have a previous specific association with DA neurons; one interesting exception is chemokine C-C motif ligand 2 (CCL2). Although originally described as a mediator of inflammatory responses, CCL2 also has an unexpectedly important role in DA cell function, in that it promotes DA cell differentiation and remains constitutively expressed in mature rodent DA cells along with its receptor, CCR2 (Edman *et al*, 2008). Furthermore, application of exogenous CCL2 increases DA release and DA cell burst firing, whereas knockout of its receptor blunts cocaine-induced sensitization (Trocello *et al*, 2011). CCL2 expression was evident in human DA neurons (Figure 2). It is possible that the upregulation of CCL2 expression (Table 2; Figure 1) represents a molecular mechanism contributing to the plasticity of DA neurons following chronic exposure to cocaine.

One of the more striking findings of the study was the significant downregulation of numerous genes associated with the midbrain DA neuron phenotype (Table 2; Figures 1 and 3b). The abundance of SLC6A3 transcript encoding the DA transporter (a primary determinant of extracellular DA levels and obligatory mediator of cocaine's rewarding effects) was diminished in cocaine abusers, as previously described (Bannon *et al*, 2002; Little *et al*, 1998; Zhou *et al*, 2014). Reduced expression of the TH gene encoding tyrosine hydroxylase (the rate-limiting enzyme in DA production) is consistent with reductions in DA levels reported in chronic cocaine abusers (Wilson *et al*, 1996). The effect of cocaine on expression of most other genes associated with DA cell phenotype has not been investigated to the same extent. Animal studies, however, demonstrate that genetic modifications reducing expression of the DA phenotype-specifying factor FOXA2 or the DA cell trophic factor receptor RET lead to decrements in DA cell number and function over time (Kittappa *et al*, 2007; Kramer *et al*, 2007); it seems plausible that the reductions in FOXA2 and RET gene expression seen in chronic cocaine abusers could render their DA cells similarly vulnerable. We also observed reduced expression of the ACHE gene encoding acetylcholinesterase, an enzyme normally expressed in abundance in human DA neurons (Landwehrmeyer *et al*, 1993). In agreement with a previous report (Meador-Woodruff *et al*, 1993), we did not observe decreased midbrain expression of DRD2 (encoding the DA autoreceptor) in cocaine abusers, although we did observe a significant downregulation of KCNJ6, an inwardly rectifying potassium channel that mediates DRD2 signaling (Inanobe *et al*, 1999). Any or all of these differences could arguably contribute to a developing dysfunction of DA neurons. Although not all of genes associated with DA phenotype were significantly influenced by cocaine abuse, at least two-thirds of the genes downregulated in our dataset can be presumed to be expressed predominantly in DA neurons, as evidenced by profound decreases in their expression in the ventral midbrain of Parkinson's disease subjects (Parkinson's disease gene expression database http://www2.cancer.ucl.ac.uk/Parkinson_Db2/; Taccioli *et al*, 2011). Overall, these data provide rather strong evidence that chronic cocaine abuse is associated with a broad (though not all-encompassing) diminution of the midbrain DA cell phenotype. Compromised functionality of DA neurons could underlie some of the adverse consequences seen with continued cocaine abuse and upon cessation of its use (Koob and Volkow, 2010), and may exacerbate decrements in DA signaling seen in humans in the course of the normal aging process (Björklund and Dunnett 2007; Dowling *et al*, 2008).

Among the downregulated genes, melanocortin 2 receptor accessory protein 2 (MRAP2; previously designated as C6orf117), encodes an accessory protein that regulates melanocortin receptor (in particular, MCR-4) signaling (Asai *et al*, 2013). MCR-4 has garnered interest in the drug abuse field because its expression in the rodent forebrain is increased by chronic cocaine administration, whereas MCR-4-null mice are resistant to some behavioral effects of cocaine (Hsu *et al*, 2005). Although MCR-4 was not detected in our samples, we found that MRAP2 was robustly expressed in human DA neurons (with an apparent nuclear localization; Figure 2) and was downregulated by chronic cocaine abuse in a ROC-significant manner (Table 2; Figure 1). These data suggest that MRAP2 may serve a novel function specific to DA neurons.

We note that some transcripts differentially expressed in cocaine abusers as a group were not statistically diagnostic in our ROC curve analysis due to the extent of inter-subject heterogeneity. This variance could arise for a myriad of reasons; as an example, we recently determined that the magnitude of HSPA1A and HSPA1B (ie, HSP70) gene expression in cocaine abusers can be influenced by certain peri-mortem events (Johnson *et al*, 2012). Perhaps more surprising was the fact that, despite inherent differences in human biology, circumstances of death, and study sample size, approximately half of all the genes differentially expressed were diagnostically predictive for assignment of subjects to the cocaine abuse or control cohort (Table 2). In aggregate, differences in the profiles of gene expression constitute a molecular signature of chronic cocaine abuse that most likely reflects some core pathophysiological processes occurring in DA neurons.

Several lines of evidence (eg, localization experiments or the loss of expression concurrent with DA cell loss in Parkinson's disease) substantiate that many of the genes most impacted in the cocaine abusers' midbrain are robustly expressed (in some cases, expressed exclusively) in DA neurons under basal conditions. That said, the present experiments did not directly examine the cellular loci of their differential expression. Astroglia and microglia can contribute to the process of drug addiction (see, for example, Reissner *et al*, 2014; Schwarz and Bilbo, 2013); it is plausible that the glia intimately associated with DA neurons may substantively contribute to differences in the profiles of the midbrain gene expression observed in cocaine abusers. Differential gene expression most likely

varies across subgroups of DA cells as well. Future studies will be needed to advance our understanding of these issues. Determining the contributions of specific changes in gene expression to the processes underlying drug addiction will ultimately require experimental interventions in model systems or man.

In previous postmortem studies (Albertson *et al*, 2004; 2006), we identified robust decreases in the forebrain expression of myelin basic protein and other myelin-related genes not anticipated by drug abuse models of the time. Subsequent imaging studies in cocaine-dependent subjects documented alterations in myelin and white matter integrity that correlated with measures of impulsivity and treatment outcome, while recent rodent studies have not only modeled these cocaine-induced changes but also shown their amelioration using a pharmacological intervention (see Kovalevich *et al*, 2012; Xu *et al*, 2010). These data support the contention that human postmortem studies represent a valuable complement to other avenues of inquiry in identifying potential biomarkers and novel therapeutic targets for the treatment of drug abuse.

## FUNDING AND DISCLOSURE

This work was supported by NIH grants DA006470 (MJB) and DA032222 (MMJ). The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

We thank Dr Rodrigo Andrade for helpful comments regarding the study, Drs Steve Krawetz and Susan Dombrowski for advice regarding bioinformatic analyses, Drs Sokol Todi and Wei-Ling Tsou for assistance with microscopy, and Ms Candace Savonen for assistance with manuscript preparation and GEO submissions.

## REFERENCES

Albertson DN, Pruetz B, Schmidt CJ, Kuhn DM, Kapatos G, Bannon MJ (2004). Gene expression profile of the nucleus accumbens of human cocaine abusers: evidence for dysregulation of myelin. *J Neurochem* **88**: 1211–1219.

Albertson DN, Schmidt CJ, Kapatos G, Bannon MJ (2006). Distinctive profiles of gene expression in the human nucleus accumbens associated with cocaine and heroin abuse. *Neuropsychopharm* **31**: 2304–2312.

Asai M, Ramachandrappa S, Joachim M, Shen Y, Zhang R, Nuthalapati N *et al* (2013). Loss of function of the melanocortin 2 receptor accessory protein 2 is associated with mammalian obesity. *Science* **341**: 275–278.

Bannon MJ, Pruetz B, Barfield E, Schmidt CJ (2004). Transcription factors specifying dopamine phenotype are decreased in cocaine users. *Neuroreport* **15**: 401–404.

Bannon MJ, Pruetz B, Manning-Bog AB, Whitty CJ, Michelhaugh SK, Sacchetti P *et al* (2002). Decreased expression of the transcription factor NURR1 in dopamine neurons of cocaine abusers. *Proc Natl Acad Sci USA* **99**: 6382–6385.

Bannon MJ, Whitty CJ (1997). Age-related and regional differences in dopamine transporter mRNA expression in human midbrain. *Neurology* **48**: 969–977.

Beaumont TL, Yao B, Shah A, Kapatos G, Joeb JA (2012). Layer-specific CREB target gene induction in human neocortical epilepsy. *J Neurosci* **32**: 14389–14401.

Berhow MT, Hiroi N, Kobierski LA, Hyman SE, Nestler EJ (1996). Influence of cocaine on the JAK–STAT pathway in the mesolimbic dopamine system. *J Neurosci* **16**: 8019–8026.

Björklund A, Dunnett SB (2007). Dopamine neuron systems in the brain: an update. *Trends Neurosci* **30**: 194–202.

Breitling R, Armengaud P, Amtmann A, Herzyk P (2004). Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett* **573**: 83–92.

DeArmond SJ, Fusco MM, Dewey MM (1989). *Structure of the Human Brain: A Photographic Atlas*. 3rd edn Oxford University Press: New York, NY, USA.

Dhar K, Rakesh K, Pankajakshan D, Agrawal DK (2013). SOCS3 promotor hypermethylation and STAT3-NF-KB interaction down-regulate SOCS3 expression in human coronary artery smooth muscle cells. *Am J Physiol Heart Circ Physiol* **304**: H776–H785.

Dowling GJ, Weiss SRB, Condon TP (2008). Drugs of abuse and the aging brain. *Neuropsychopharm* **33**: 209–218.

Edman LC, Mira H, Arenas E (2008). The β-chemokines CCL2 and CCL7 are two novel differentiation factors for midbrain dopaminergic precursors and neurons. *Exp Cell Res* **314**: 2123–2130.

Feng J, Nestler EJ (2013). Epigenetic mechanisms of drug addiction. *Curr Opin Neurobiol* **23**: 521–528.

Franks TM, Lykke-Andersen J (2007). TTP and BRF proteins nucleate processing body formation to silence mRNAs with AU-rich elements. *Genes Dev* **21**: 719–735.

Gingras S, Pelletier S, Boyd K, Ihle JN (2007). Characterization of a family of novel cysteine- serine-rich nuclear proteins (CSRNP). *PLoS ONE* **28**: e808.

Hsu R, Taylor JR, Newton SS, Alvaro JD, Haile C, Han G *et al* (2005). Blockade of melanocortin transmission inhibits cocaine reward. *Eur J Neurosci* **21**: 2233–2242.

Hurt E, Luo M-J, Rother S, Reed R, Straßer K (2004). Cotranscriptional recruitment of the serine-arginine-rich (SR)-like proteins Gbp2 and Hrb1 to nascent mRNA via the TREX complex. *Proc Natl Acad Sci USA* **101**: 1858–1862.

Inanobe A, Yoshimoto Y, Horio Y, Morishige K-I, Hibino H, Matsumoto T *et al* (1999). Characterization of G-protein-gated $K^+$ channels composed of Kir3.2 subunits in dopaminergic neurons of the substantia nigra. *J Neurosci* **19**: 1006–1017.

Johnson MM, David JA, Michelhaugh SK, Schmidt CJ, Bannon MJ (2012). Increased heat shock protein 70 gene expression in the brains of cocaine-related fatalities may be reflective of postdrug survival and intervention rather than excited delirium. *J Forensic Sci* **57**: 1519–1523.

Johnson MM, Michelhaugh SK, Bouhamdan M, Schmidt CJ, Bannon MJ (2011). The transcription factor NURR1 exerts concentration-dependent effects on target genes mediating distinct biological processes. *Front Neurosci* **5**: 135.

Karch SB (2002). *Karch's Pathology of Drug Abuse*. 3rd edn CRC Press: Boca Raton, FL, USA.

Kittappa R, Chang WW, Awatramani RB, McKay RDG (2007). The foxa2 gene controls the birth and spontaneous degeneration of dopamine neurons in old age. *PLoS Biol* **512**: 2875–2884.

Koob GF, Volkow ND (2010). Neurocircuitry of addiction. *Neuropsychopharm* **35**: 217–238.

Kovalevich J, Corley G, Yen W, Rawls SM, Langford D (2012). Cocaine-induced loss of white matter proteins in the adult mouse nucleus accumbens is attenuated by administration of a β-lactam antibiotic during cocaine withdrawal. *Am J Pathol* **181**: 1921–1927.

Kramer ER, Aron L, Ramakers GM, Seitz S, Zhuang X, Beyer K *et al* (2007). Absence of Ret signaling in mice causes progressive and late degeneration of the nigrostriatal system. *PLoS Biol* **53**: 0616–0628.

Kufahl PR, Zavala AR, Singh A, Thiel KJ, Dickey ED, Joyce JN *et al* (2009). c-Fos expression associated with reinstatement of cocaine-seeking behavior by response-contingent conditioned cues. *Synapse* **63**: 823–835.

Landwehrmeyer B, Probst A, Palacios JM, Mengod G (1993). Expression of acetylcholinesterase messenger RNA in human brain: an in situ hybridization study. *Neurosci* **57**: 615–634.

Little KY, McLaughlin DP, Zhang L, McFinton PR, Dalack GW, Cook EH Jr *et al* (1998). Brain dopamine transporter messenger RNA and binding sites in cocaine users: a postmortem study. *Arch Gen Psychiatry* **55**: 793–799.

Ma DK, Jang M-H, Guo JU, Kitabatake Y, Chang M-L, Powanpongkul N *et al* (2009). Neuronal activity–induced Gadd45b promotes epigenetic DNA demethylation and adult neurogenesis. *Science* **323**: 1074–1077.

McCullumsmith RE, Hammond JH, Shan D, Meador-Woodruff JH (2013). Postmortem brain: An underutilized substrate for studying severe mental illness. *Neuropsychopharmacology* **39**: 65–87.

McKittrick E, Gafken PR, Ahmad K, Henikoff S (2004). Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proc Natl Acad Sci USA* **101**: 1525–1530.

Meador-Woodruff JH, Little KY, Damask SP, Mansour A, Watson SJ (1993). Effects of cocaine on dopamine receptor gene expression: A study in the postmortem human brain. *Biol Psychiatry* **34**: 348–355.

Metz CE (2006). Receiver operating characteristic analysis: A tool for the quantitative evaluation of observer performance and imaging systems. *J Am Coll Radiol* **3**: 413–422.

Moratalla R, Elibol B, Vallejo M, Graybiel AM (1996). Network-level changes in expression of inducible Fos-Jun proteins in the striatum during chronic cocaine treatment and withdrawal. *Neuron* **17**: 147–156.

Ploski JE, Newton SS, Duman RS (2006). Electroconvulsive seizure-induced gene expression profile of the hippocampus dentate gyrus granule cell layer. *J Neurochem* **99**: 1122–1132.

Reissner KJ, Brown RB, Spencer S, Tran PK, Thomas CA, Kalivas PW (2014). Chronic administration of the methylxanthine propentofylline impairs reinstatement to cocaine by a GLT-1-dependent mechanism. *Neuropsychopharmacology* **39**: 499–506.

Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassman M *et al* (2006). The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol* **7**: 3.

Schwarz JM, Bilbo SD (2013). Adolescent morphine exposure affects long-term microglial function and later-life relapse liability in a model of addiction. *J Neurosci* **33**: 961–971.

Stan AD, Ghose S, Gao XM, Roberts R, Lewis-Amezcua K, Hatanpaa KJ *et al* (2006). Human postmortem tissue: What quality markers matter? *Brain Res* **1123**: 1–11.

Taccioli C, Maselli V, Tegnér J, Gomez-Cabrero D, Altobelli G, Emmett W *et al* (2011). ParkDB: a Parkinson's disease gene expression database. *Database* **2011**: bar007.

Thomas KL, Arroyo M, Everitt BJ (2003). Induction of the learning and plasticity-associated gene Zif268 following exposure to a discrete cocaine-associated stimulus. *EurJNeurosci* **17**: 1964–1972.

Tong Q, Tsai J, Tan G, Dalgin G, Hotamisligil GS (2005). Interaction between GATA and C/EBP family of transcription factors is critical in GATA-mediated suppression of adipocyte differentiation. *Mol Cell Biol* **25**: 706–715.

Trocello JM, Rostene W, Melik-Parsadaniantz S, Godefroy D, Roze E, Kitabgi P *et al* (2011). Implication of CCR2 chemokine receptor in cocaine-induced sensitization. *J Mol Neurosci* **44**: 147–151.

Volkow ND, Wang GJ, Fowler JS, Tomasi D, Telang F (2011). Addiction: beyond dopamine reward circuitry. *Proc Natl Acad Sci USA* **108**: 15037–15042.

Whitty CJ, Paul MA, Bannon MJ (1997). Neurokinin receptor mRNA localization in human midbrain dopamine neurons. *J Comp Neurol* **382**: 394–400.

Wilson JM, Levey AI, Bergeron C, Kalasinsky K, Ang L, Peretti F *et al* (1996). Striatal dopamine, dopamine transporter, and vesicular monoamine transporter in chronic cocaine users. *Ann Neurol* **40**: 428–439.

Xu J, DeVito EE, Worhunsky PD, Carroll KM, Rounsaville BJ, Potenza MN (2010). White matter integrity is associated with treatment outcome measures in cocaine dependence. *Neuropsychopharmacology* **35**: 1541–1549.

Zhou Y, Michelhaugh SK, Schmidt CJ, Liu JS, Bannon MJ, Lin Z (2014). Ventral midbrain correlation between genetic variation and expression of the dopamine transporter gene in cocaine-abusing versus non-abusing subjects. *Addict Biol* **19**: 122–131.



Supplementary Information accompanies the paper on the Neuropsychopharmacology website (http://www.nature.com/npp)