Gene Function Analysis of Biological Processes Affected by Fentanyl Use

Srinivasa Iyer

Vista del Lago High School, Folsom, California

Abstract

The opioid epidemic, fueled by increased use of opioids like fentanyl, claims close to 50,000 deaths annually, out of about ten million users. However, the genetic aspects of fentanyl consumption are relatively understudied and a greater understanding of genetic biomarkers of fentanyl usage would be immensely beneficial to the public health system. This study aimed to analyze post-mortem brain tissue samples of fentanyl-using subjects and healthy controls to identify biomarkers of fentanyl use. 186 differentially expressed genes were selected and entered into a protein interaction network, following which commonly present functions and pathways were analyzed. Processes including inflammatory response and responses to the metabolic process were identified as potential biomarkers. This study highlighted the connections between opioid use and the immune system, an emerging area of study.

Introduction

The rise of opioid usage in the United States has led to a public health emergency, with over ten million people using opioids annually [1]. This has led to an unprecedented number of overdose deaths, numbering over 48,000 in 2023 [1]. Fentanyl overdoses factor in over 50% of all opioid overdose fatalities and is one of the few opioids that has recorded an increase in consumption in the last decade [2][3].

Fentanyl is a semi-synthetic opioid that binds to plasma proteins in human fat upon ingestion [3]. It is then transported by the blood stream to the brain, where it preferentially acts upon the μ-opioid receptors, and to a lesser extent, the κ-opioid receptors, and the delta opioid receptors [3]. Upon reception, these receptors produce a state of euphoria, following which the opioid is metabolized by the liver and then excreted [3].

Unregulated consumption of greater than controlled (> 0.00005 grams) quantities of fentanyl, leading to overdose is one of the leading causes for opioid-related deaths in the United States [4]. Gene expression changes caused by fentanyl usage is an understudied area due to the recentness of the fentanyl epidemic.

Gene expression analysis involves analyzing changes in messenger RNA (mRNA) expression between healthy control samples and the target samples [5]. mRNA is synthesized from instructions provided by deoxyribonucleic acid (DNA) [5]. Research has shown that minor changes in mRNA expression could lead to major phenotypic differences [6]. Gene analysis has been instrumental in advancing our understanding of diseases and their treatments: for example,

gene analysis has been instrumental in understanding the disease progression, symptom severity and potential drug targets for breast cancer [7][8][9].

Multiple research studies have been conducted to understand the effects of fentanyl usage on cellular function of mice (*Mus musculus*). A study found that perinatal exposure to fentanyl to female mice led to them having smaller litters compared to controls [10]. Perinatal exposure to opioids like fentanyl also led to differentiated development of the offspring, lasting at least till adolescence [11]. Fentanyl is used in controlled quantities (> 0.00005 grams) as a painkiller in medical settings [3]. Research on gene expression of human brain cells in such settings has shown that expression of genes involved with synaptic transmission, inflammation, and organization of the extracellular matrix changed significantly [12].

Understanding the gene expression changes wrought by fentanyl consumption to identify potential methods to counteract the effects of fentanyl in urgent cases would go a long way in saving human lives.

Materials and Methods

GEO and GEO2R

Gene Expression Omnibus (GEO), an online National Center for Biotechnology Information (NCBI) repository containing public gene expression data from various studies (Fig.1) [13]. The database was searched using the terms "opioid abuse", "fentanyl", "expression profiling by array" and "*Homo sapiens*", following which dataset GSE174409 was selected. The dataset contains post-mortem brain samples of healthy controls and of opioid users [14]. Drugs detected among opioid users include fentanyl, heroin, cocaine, and morphine among others [14]. The studies used microarrays for data collection. Microarrays are a chip-based technology that uses oligonucleotide probes to capture strands of complementary DNA (cDNA). The cDNA is reverse transcribed from the mRNA present in the brain samples [15].

The GEO-provided tool GEO2R was used to conduct a differential expression analysis study of healthy controls (Group 1) versus patient samples where fentanyl was detected (Group 2) [15]. GEO2R processes gene expression values and outputs a table of differentially expressed genes (DEGs) between multiple groups. Several thousand genes were deemed statistically significant by GEO2R for each dataset.

GEO2R also calculated the fold changes for each gene based on differential expression between Group 1 and Group 2. Fold change is a ratio of the average expression value of a gene in one group divided by the average expression value in a different group: a fold change greater than one indicates overexpression while a fold change lesser than 1 indicates underexpression [16].

GEO2R was also used to verify a normal distribution of gene expression values, and no outliers were present. After the lists of DEGs were obtained, they were processed in Excel. The top 250 most differentially expressed genes, sorted by increasing *p*-values, were retained in the list and the others were deleted.

STRING

The final list of the top 250 most differentially expressed genes were entered into the Search Tool for the Retrieval of Interacting Proteins (STRING) tool, out of which 186 genes were recognized. The 186 genes were then used to create a network of known and predicted protein-protein interactions of the gene products. The STRING creates a network of known and predicted protein-protein interactions by scanning various databases (Fig.1) [17]. It also lists related Gene Ontology (GO) processes and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Fig. 1).
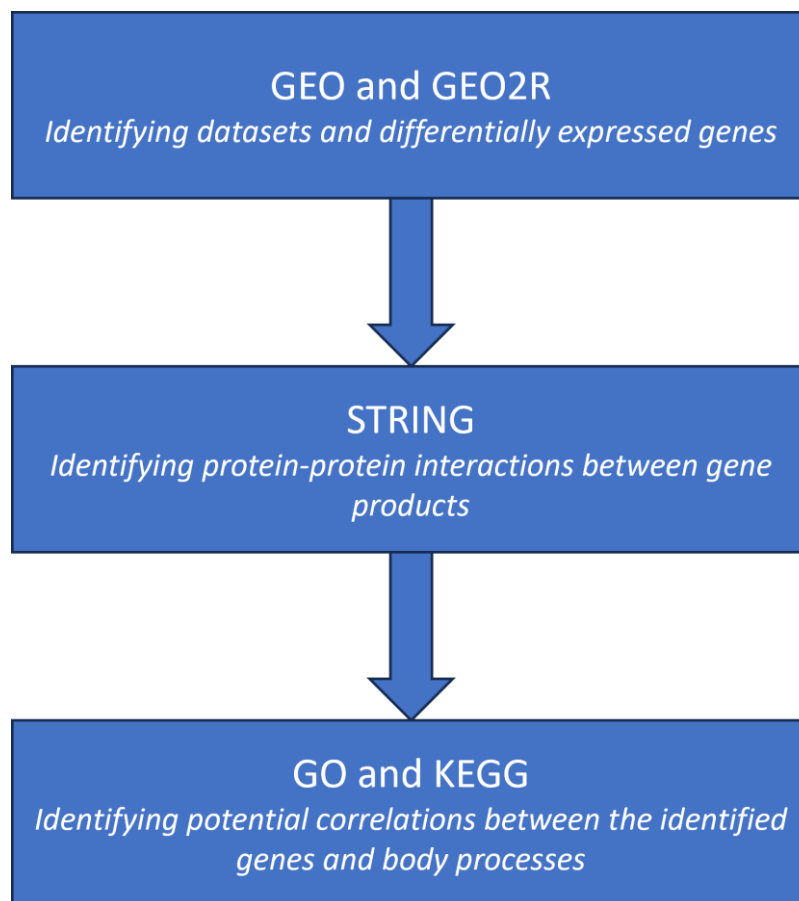

Figure 1: A flowchart of the process used in the experiment to analyze the genes and gene products.

Results and Discussion

Fig. 2 visualizes the fold change and the adjusted *p*-value of genes in affected samples compared to healthy controls. Fold change (FC) represents the difference between the expression of the gene in healthy control samples compared to the target samples [16].
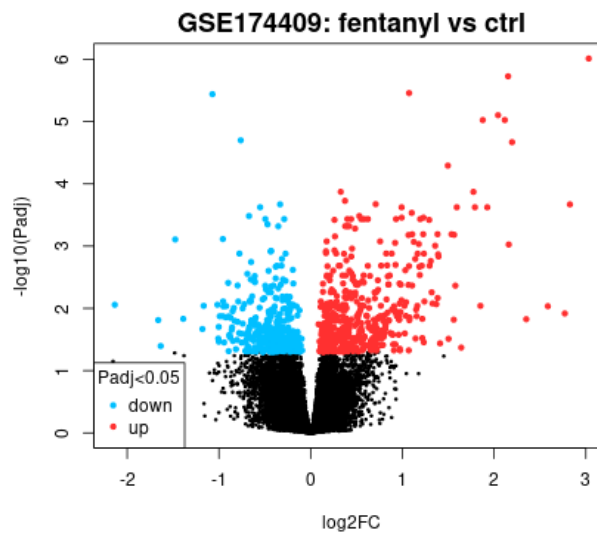

Figure 2: Volcano plot showing the fold change and adjusted p-*values of the genes in the dataset.*

Positive values indicate an upregulation of the gene while negative values indicate downregulation. The chart normalizes the exponentially different FC values through logarithms (logFC) [16]. LogFC values represent the relative upregulation or downregulation of the gene in targets samples versus control samples in a easily comprehensible scale [16].

 p-values represent the probability of the hypothesis occurring due to chance: therefore, lower *p*-values indicate greater statistical significance [18]. The p-value was set to an upper bound 0.05 in order to highlight genes with greater differentiation in the their expression between controls and affected samples yields a lower *p*-value. The negative logarithm of the *p*-value, as shown in the figure, serves to normalize the exponentially different *p*-values, into a more comprehensible scale for humans. The top 250 genes with the smallest *p*-values were selected for further analysis.
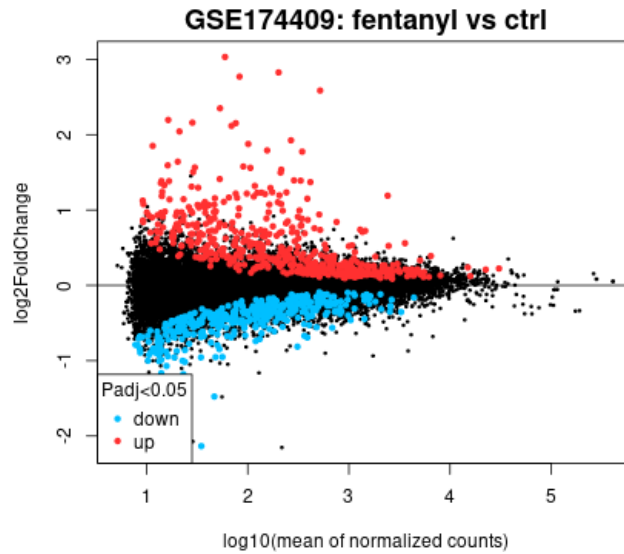
Figure 3: Bland-Altman plot visualizing the relative upregulation or downregulation compared to the overall mean expression

Figure 3 is a Bland-Altman plot that measures the agreement between the difference between two expressions on the *y*-axis and the mean of all samples of both expressions on the *x*-axis [19]. In this case, the Bland-Altman plot contains the fold change on the *y*-axis and the mean expression of the gene across all samples on the *x*-axis. Both the fold change and the mean gene expression are modified through logarithms to assume a linear scale for an increased accuracy.
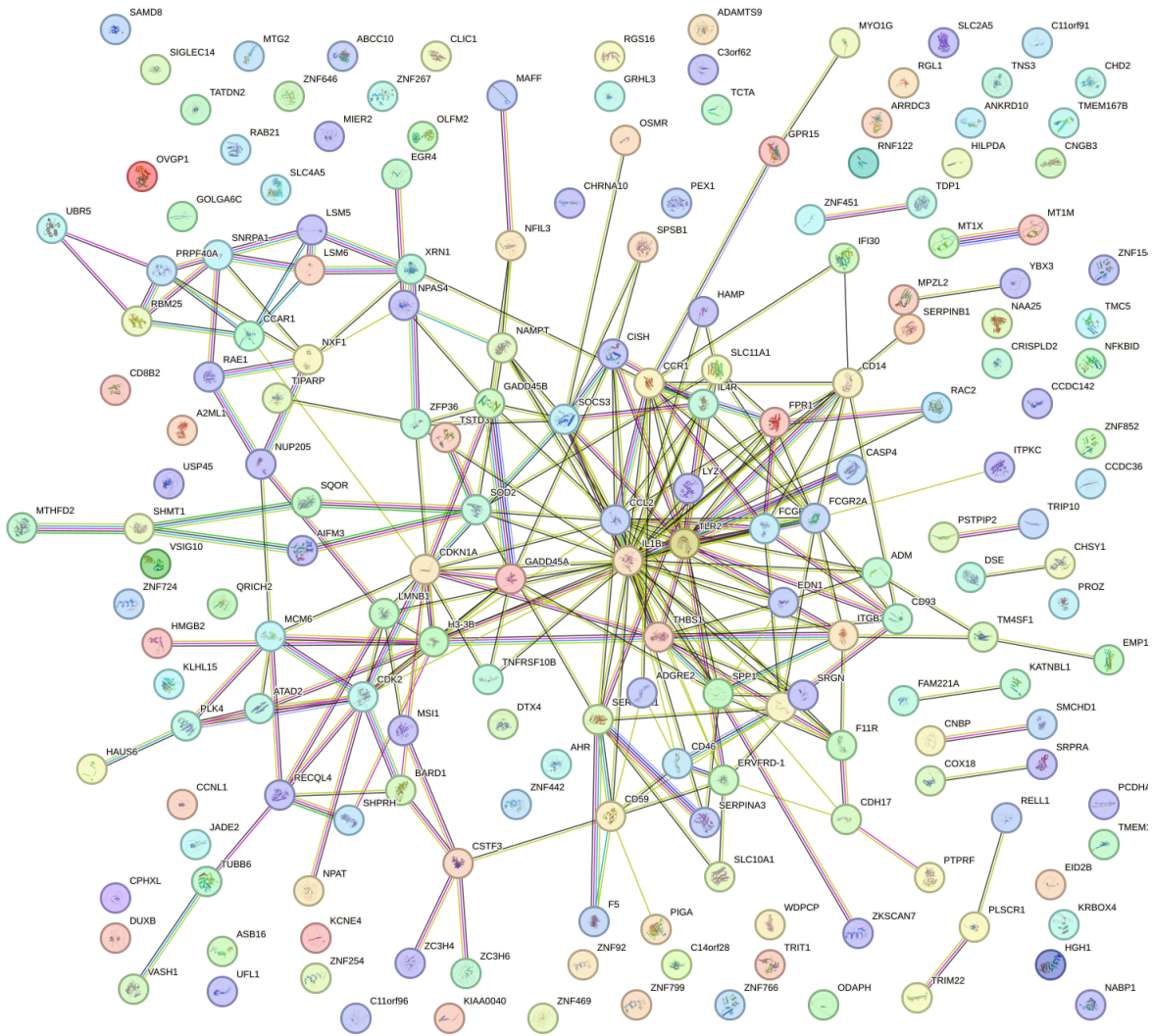
Figure 4: The STRING network of 186 most significant genes.

The top 250 most differentially expressed genes, determined through their *p*-values, were submitted to the STRING tool. The *p*-value cutoff was set to 0.05.

The figure represents the STRING network of the top 250 most differentially expressed genes, determined by their *p*-values and logFC.

The 250 genes that passed all cut-off criteria were submitted to the STRING tool, out of which 186 were successfully mapped. The minimum confidence score for whether a protein interaction existed was set to 0.050. The protein-protein interaction (PPI) *p*-value was 1.1×10^{-16}, indicating that the network had significantly more interactions than expected (252 connections instead of an expected 143).
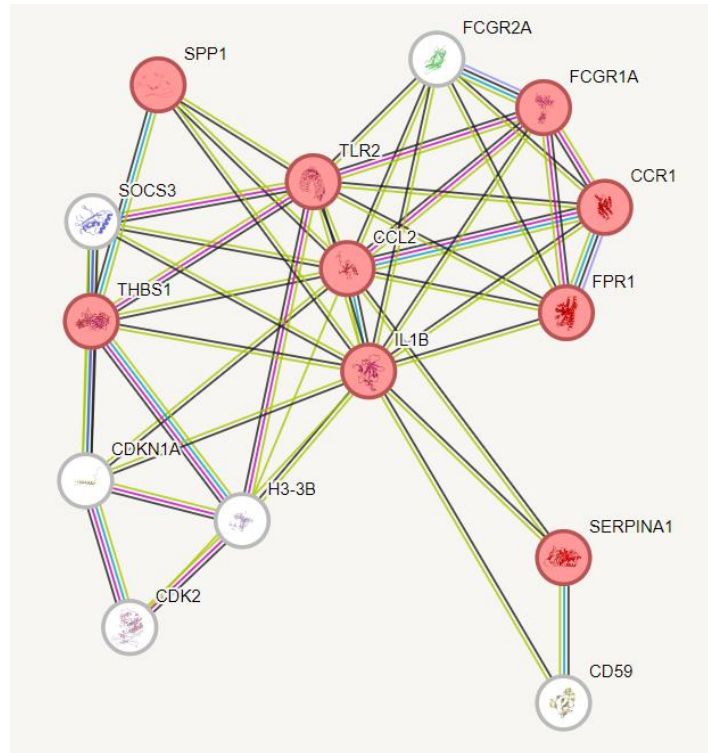
Figure 5: Interactions between the genes with maximum interactions in the original network (Fig. 4). The genes that code for products involved with inflammatory response highlighted in red.

The top 15 genes with the maximum protein interactions, determined by their node degrees, were then separated into a separate network (Fig. 5) where the *p*-value was set to 0.05. The node degree of a gene is the number of interactions it has with other gene products in the network [20]. The following table lists the node degrees of the top 15 genes.

Gene	Node Degree
IL1B	38
CCL2	21
TLR2	18
CDKN1A	14
CCR1	13
FCGR2A	13
H3-3B	13
SOCS3	12
THBS1	12
CDK2	11
SPP1	11
CD59	9
FCGR1A	9
FPR1	9

Of the top 15 genes, 8 genes were connected to the function protein-containing complex binding, a molecular function where enzymes attach to macromolecular complexes [21]. Additionally, 22 of the top 186 genes in the first network were connected to inflammatory response. Inflammatory response is the immune system's response to flush pathogens and toxins from the blood steam [22]. The correlation between the drugs and the immune system's response to it reflects the immune system's involvement in actively combating the presence of drug-related toxins in the blood. The significant upregulation of genes relating to inflammatory response highlights the fact that the body responds to a higher amount of toxins during drug consumption. The inability of the body's inflammatory response to combat increased drug consumption could potentially lead to an overdose and death.

Also, several functions related to the macromolecule metabolic process were enriched in the STRING network. Genes involved in such processes produce gene products that modulate the frequency, rate or extent of the chemical reactions and pathways involving macromolecules [23]. Out of the 186 genes in the first network, 49 gene products negatively regulated the macromolecule metabolic process while 91 genes products were involved in some form of regulating the said process. This reinforces the fact that drug consumption fundamentally changes human gene expression and fentanyl exposure leads to significant downregulation of vital gene products relating to metabolism.

The effects of drug abuse are therefore not simply psychological, but also affect the fundamental biology of humans through the regulation of vital processes like inflammatory response and the metabolic processes. Further research into the biological basis of drug abuse is required to ascertain if the effects of changed gene expression are severe. Also, these correlations could also be used in the future to identify potential drug targets to overcome opioid consumption-related consequences.

Acknowledgements

References

1. https://drugabusestatistics.org/opioid-epidemic/
2. https://health-policy-systems.biomedcentral.com/articles/10.1186/s12961-020-00596-8
3. https://drugabusestatistics.org/fentanyl-abuse-statistics/
4. https://nida.nih.gov/research-topics/trends-statistics/overdose-death-rates
5. https://www.ncbi.nlm.nih.gov/probe/docs/applexpression/
6. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6346310/
7. https://www.science.org/doi/10.1126/science.1090005

8. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4790538/
9. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5575784/
10. https://pubmed.ncbi.nlm.nih.gov/19088018/
11. https://onlinelibrary.wiley.com/doi/10.1111/adb.12895
12. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8051687/
13. https://www.ncbi.nlm.nih.gov/geo/info/geo2r.html
14. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE174409
15. https://www.nature.com/scitable/definition/microarray-202/
16. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7230499/
17. https://string-db.org/cgi/about
18. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4111019/
19. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4470095/
20. https://www.sciencedirect.com/science/article/abs/pii/B9780124079083000042
21. https://amigo.geneontology.org/amigo/term/GO:0044877
22. https://amigo.geneontology.org/amigo/term/GO:0006954
23. https://amigo.geneontology.org/amigo/term/GO:0060255