Transcriptomic Analysis of the Receptors Implicated in NeuroAIDS and Establishing *in-Silico-Ayurvedic* Remedy of for NeuroAIDS

Bhagyashri Manohar Jadhao[1*], Kruthi. KT[1], Manasi Bhoi[1], Shylesh Murthy IA[2] and Preenon Bagchi[3]
[*]Padmashree Institute of Management and Sciences, Bengaluru, India.
[2]Vasishth Academy of Advanced Studies and Research, Bengaluru, India.
[3]Institute of Biosciences and Technology, MGM University, Chh. Sambhajinagar, India.
Corresponding author: Email: beingbmj1996@gmail.com

Abstract

The human immunodeficiency virus (HIV) is the cause of acquired immunodeficiency syndrome (AIDS), a chronic illness that may be fatal. HIV weakens our immune system, which makes it difficult for our body to fight against infection and disease. This is a potentially fatal condition. HIV, also known as the human immunodeficiency virus, is a virus that attacks immune-supporting cells in the body, making the person more vulnerable to other infections and diseases. HIV can spread through sharing injecting supplies, having sex with an infected individual, or engaging in unprotected sex (sex without the use of a condom or HIV medicine to prevent or treat HIV). If therapy is not received, HIV can lead to AIDS (acquired immunodeficiency syndrome). HIV cannot be cured and cannot be removed from the human body. Research is being done to determine the cause of the changes that an HIV infection causes to the central nervous system (CNS). HIV-associated dementia (HAD; also known as AIDS dementia and HIV encephalopathy) and neurocognitive impairment are two of the neurological consequences of AIDS (NeuroAIDS). The most serious and debilitating HIV-related central nervous system (CNS) problem is hypoxic encephalopathy (HAD). The pathophysiology, clinical characteristics, and neurobiological elements of HAD have been better understood recently, but research into these areas is still quite difficult. The interpretation of the diagnostic aspects of HIV-associated minor cognitive/motor disorder (MCMD), a milder form of HAD, and HIV neuroinvasion requires an understanding of the mechanisms of CNS proliferation, HIV neuroinvasion, and HAD pathogenesis, respectively. Looking into our genes, transcriptomic analysis shows us which receptors are hit by neuroAIDS and what they do in the illness. It points out important genes and biological pathways that play a part, making it easier to spot targets for treatment. By mapping out individual gene patterns, we can tailor treatments just right, picking up on unique markers for each person. This information is key for cooking up digital Ayurvedic strategies against neuroAIDS, pushing forward medicines aimed just where they're needed, and getting to the bottom of why this disease happens at a molecular level.

Keywords: *neuroAIDS, transcriptomics, RNA-SEQ, next-generation sequencing, Ayurvedic medicinal herbs, modelling, Lipinski's rule of five, docking*

1. Introduction

As Human Immunodeficiency Virus infection (HIV) and Acquired Immune Deficiency Syndrome (AIDS) combat with our immune system, our central nervous system also gets affected in the condition called Neurological Acquired Immune Deficiency Syndrome (neuroAIDS) (Kauletal., 2005; McArthur, 1987; McArthur, 1993). In this condition, there are neurological complications along with the causal factor of HIV infection (Kranick, & Nath, 2012). This disorder is galloped around dysfunction and degeneration of CNS neurons that leads to a number of clinical syndromes involving behavioural, cognitive, and motor functions wherein abnormal potentials, other electrophysiological responses are evoked, and this disease incorporates both infectious and degenerative pathophysiologic pathways (Felger & Treadway, 2017; Fernandez-Cruz & Fellows, 2017; Shapshak etal., 2011). Interlocking neurologic and psychiatric (neuropsychiatric) disorders account for more than 15% of the global illness burden; these syndromes include mood disorders, schizophrenia, addiction, dementia, epilepsy, and chronic pain disorders (Minagar etal., 2008; McCombe etal., 2009; Ances, 2008). When a huge worldwide infectious epidemic and nervous system disease intersect, it can lead to neuropsychiatric diseases linked to HIV infection (Minagar et al., 2008; McCombe etal., 2009; Ances, 2008). This can result in significant morbidity and mortality, particularly when the condition worsens and becomes AIDS (Minagar et.

al., 2008; McCombe etal., 2009; Ances, 2008).

A wide range of neuropsychiatric disorders, collectively referred to as neuroAIDS, are caused by HIV's direct infection and damage to the central and peripheral nervous systems (Kranick, & Nath, 2012). These disorders include neurocognitive disorders like HIV-associated dementia, minor neurocognitive disorder, mania/psychosis, anxiety, depression, seizures, and myelopathy and neuropathy, which are accompanied by chronic neuropathic pain and physical disabilities (Minagar et al., 2008; McCombe etal., 2009; Ances, 2008; Kranick, & Nath, 2012). In actuality, HIV enters the nervous system early after initial infection (neuroinvasion) but continues to do so because it causes glial cell chronic infection (neurotropism), which increases the risk of developing nervous system diseases (neurovirulence) (Minagar et al., 2008; McCombe etal., 2009; Ances, 2008; Kranick and Nath, 2012). Nonetheless, only a small percentage of HIV-positive individuals have a nervous system illness, suggesting that HIV causes a neuropsychiatric phenotype known as neurosusceptibility, which is characterized by age, the degree of concurrent immunosuppression, comorbidities, and genetic differences between the virus and the host (Minagar et al., 2008; McCombe etal., 2009; Ances, 2008; Kranick & Nath, 2012).

Reduced survival, higher health care expenses, and a lower quality of life are all linked to certain neuropsychiatric illnesses (Minagar et al., 2008). Children with HIV who also have developmental delays exhibit a variety of these problems. Furthermore, systemic immunosuppression raises the risk of opportunistic nervous system processes, especially in adults (Kaul etal., 2005; McCombe etal., 2009; Ances, 2008; McArthur, 1993). These include toxoplasmic encephalitis, progressive multifocal leukoencephalopathy, meningitis caused by cryptococcal and tuberculous bacteria, and primary central nervous system lymphoma, which can cause seizures, cognitive and physical impairments, psychosis, and mood disorders (Minagar et al., 2008; McCombe etal., 2009; Ances, 2008; Kranick and Nath, 2012; Kaul etal., 2005; McArthur, 1987; McArthur, 1993). HIV-1 infection causes neuropathogensis when it penetrates the central nervous system (CNS), which can lead to a variety of neurological issues both directly and indirectly. HIV infection can lead to neurological problems, one of which is AIDS dementia complex (ADC). Prior to the development of highly active antiretroviral treatment (HAART), a considerable proportion of HIV-positive patient deaths were attributed to ADC (Minagar et al., 2008; McCombe etal., 2009; Ances, 2008; Kranick, & Nath, 2012; Kaul et.al., 2005; McArthur, 1987; McArthur, 1993). This is typically a neurological symptom of HIV infection that appears in the later stages of AIDS. In the latter stages of AIDS, 30–60% of individuals with HIV-1 infection experience HIV-1-associated encephalitis, which is a brain swelling brought on by the virus that results in memory and attention issues.

Furthermore, HIV-1-associated encephalitis represents the sole brain manifestation of HIV/AIDS in 3-5% of individuals (Minagar et al., 2008; McCombe et al., 2009; Ances, 2008; Kranick, & Nath, 2012; Kaul et al., 2005; McArthur, 1987; McArthur, 1993). Numerous studies have elucidated the precise mechanism by which HIV infection might lead to neurological problems. The idea that microglia and macrophages serve as cellular reservoirs for HIV-1 infection that is productive in the brain has received support from a number of studies (Minagar et al., 2008; Kaul etal., 2005; McArthur, 1987; McArthur, 1993). Furthermore, studies have shown that HIV-1 can infect brain astrocytes and that the expression of regulatory genes such as nef and rev can indicate the presence of this infection.

Furthermore, research has demonstrated that HIV-1 can infect cells that do not express CD4, such as human skin fibroblasts, human trophoblasts, brain-derived glial and neuronal cells, colonic epithelial cells, follicular dendritic cells, muscle cells, and fetal adrenal cells (Minagar et al., 2008, McCombe et al., 2009, Ances, 2008, Kranick, & Nath, 2012; Kaul et al., 2005; McArthur, 1987; McArthur, 1993). In addition, the V3 loop of gp120 controls HIV-1 cell tropism. The infectivity of SK-N-MC cells, a CD4-negative, galactosylceramide-positive neuroblastoma cell line, is thought to be mostly determined by this loop (Ances, 2008; Kranick, & Nath, 2012; Kaul et al., 2005; McArthur, 1987).

The collective evidence points to interactions between the viral epitopes (V3 loop) and receptors as the cause of neurotropism, a nervous system cell's sensitivity to HIV-1 infection. Furthermore, HIV-1 cell tropism is regulated by the V3 loop of gp120 (Ances, 2008; Kranick, & Nath, 2012; Kaul et al., 2005; McArthur, 1987). This loop is assumed to have a major role in the infectivity of SK-N-MC cells, a CD4-negative,

galactosylceramide-positive neuroblastoma cell line. According to a body of research, neurotropism—a nervous system cell's susceptibility to HIV-1 infection—is caused by interactions between receptors and viral epitopes (the V3 loop) (Ances, 2008; Kranick, & Nath, 2012; Kaul et al., 2005; McArthur, 1987). Together with the chemokine receptors CXCR4 and CCR5, CD4 is a major receptor for HIV infection (Alkhatib, 2009; Wilen, Tilton, & Doms., 2012). Next-generation sequencing (NGS) technology has revolutionized the study of genetics linked to disease, making the creation of raw data for genome-scale investigations easier (Blankenberg & Hillman, 2014). Galaxy is a widely available tool for analyzing NGS data, and transcriptome reconstruction helps understand the genes expressed in an organism (Blankenberg, & Hillman, 2014). Transcriptomics technology studies an organism's transcriptome, which includes all RNA transcripts expressed in a single cell (Batut et al., 2018). This includes RNA levels of transcription and expression, functions, locations, trafficking, and elimination. Transcriptomics also covers splicing patterns, 5′ and 3′ end sequences, posttranscriptional modifications, and parent genes (Bradshaw, & Stahl, 2015). All transcripts, including messenger RNAs (mRNAs), microRNAs (miRNAs), and various forms of long noncoding RNAs (lncRNAs), are included in this field (Lowe et al., 2017). High-throughput sequencing techniques are used in modern transcriptomics to examine the expression of many transcripts under various physiological or pathological conditions (Bradshaw, & Stahl, 2015; Lowe et al., 2017). Additionally, computer-aided drug design has been used to establish novel ligands for neuroAIDS from Ayurvedic medicinal herbs.

2. Materials and Methods

Transcriptome Reconstruction with RNA-Seq

The RNA-seq study of neuroaids is done using the Galaxy tutorial "De novo transcriptome reconstruction with RNA-Seq" by Mallory Freeberg and Mo Heydarian, which is accessible at Galaxy Training Materials. The SRA database provided the NeuroAIDS RNA-seq data SRR1852861, SRR1852881, SRR1852860, and SRR1852859. Using the FASTQC tool (Andrews, 2000), we evaluated the read quality both before and after trimming. To improve mapping performance, low-quality bases were removed from the reads, and any repeated or overlapping reads were removed from the sequences using the Trimmomatic program (Bolger, Lohse, & Usadel, 2014). Using the HISAT2 tool, we align, or "map," the reads to the reference human genome in order to interpret the reads and their locations within the genome (Kim, Langmead, & Salzberg, 2015).

Moreover, we identify the transcript structures that the matched reads reflect by performing de novo transcriptome reconstruction. This impartial methodology allows for a thorough identification of every transcript found in our sequence. The de novo transcriptome reconstruction approach provides complete transcriptome(s) identification from the experimental samples, even if common gene/transcript databases are fairly vast but incomplete. The Stringtie tool (Kovaka et al., 2019; Pertea et al., 2015; Pertea et al., 2016) is used for this. Next, we must count the number of reads in each transcript, or count the reads based on genomic characteristics, in order to compare the abundance of transcripts across various cellular states. The tool FeatureCounts (Liao, Smyth, & Shi, 2013) is used for this.

Lastly, we estimate transcript expression using read counts or conduct differential gene expression (DGE) testing. Accurate findings can only be obtained by doing this. DESeq2 (Love, Huber, & Anders, 2014) is the tool we use for DGE. We feed it the output of FeatureCounts as input, and it applies size factor normalization, which involves computing the geometric mean of read counts for each gene across all samples, dividing each gene count by the geometric mean, and using the median of these ratios as the size factor for normalizing our sequences. In addition, we use the volcano plot (Li, Freudenberg, Suh, & Yang, 2014), a kind of scatterplot, to determine which genes are expressed.

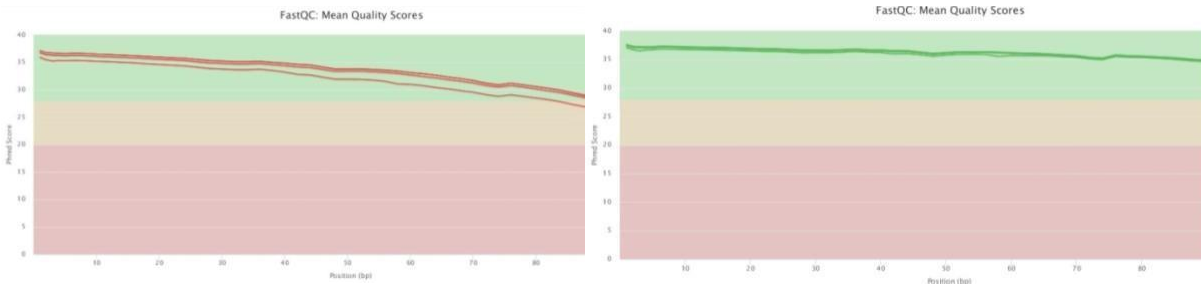
Next, we used SWISS- MODEL to model the three-dimensional structure of the receptor genes found in neuroAIDS (Waterhouse et al., 2018). Downloads of phyto-compounds were made via PUBCHEM. Phytocompounds were chosen for docking based on the concepts of Lipinski's rule of five and molinspiration software. By using docking, we are able to predict the gene receptors' preferred orientation with the phyco-

Proceedings of RSU International Research Conference (RSUCON-2024)

compounds when they bind together to form a stable complex. Patchdock was used to execute the docking process (Schneidman-Duhovny, Inbar, Nussinov, & Wolfson, 2005).

3. Results and Discussion

The Fastq sequences of the genome identified in neuroAIDS are SRR1852861, SRR1852881, SRR1852860, and SRR1852859. Sequence quality before and after trimming is noted using the FASTQC tool, and the results of FASTQC are merged using the MULTIQC tool (Ewels, Magnusson, Lundin, & Käller, 2016) for better visualization, as given in Fig. 1.



(a): MULTQC results before trimming (b): MULTQC results after trimming

Figure 1 MultiQC results before and after trimming

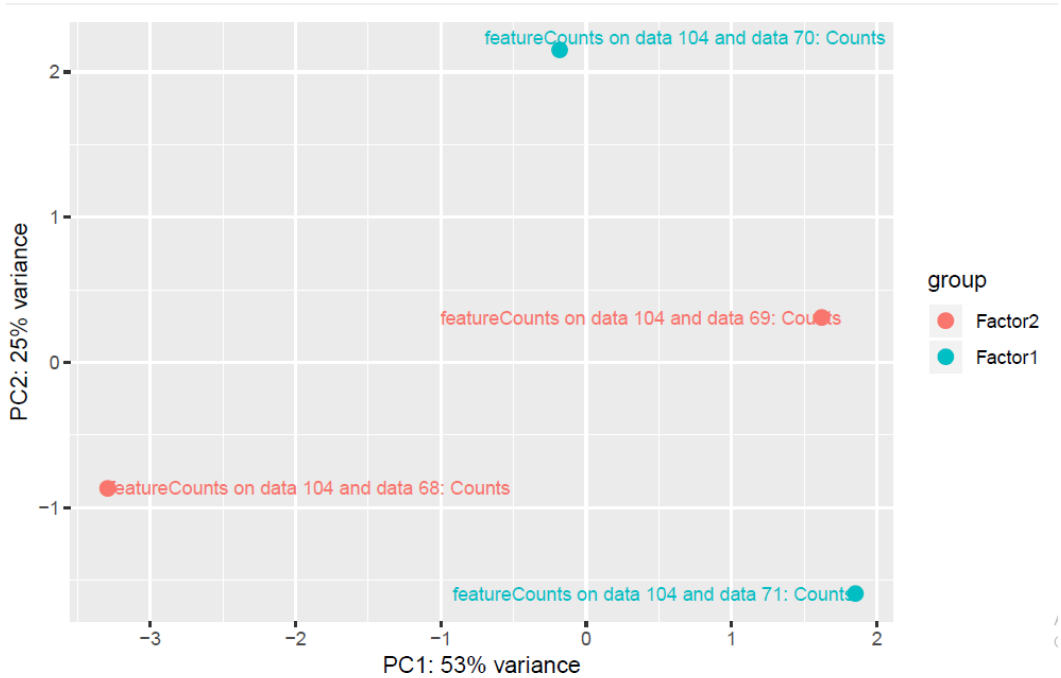
Figure 1(b) illustrates how the Trimmomatic tool trimmed or eliminated low-quality bases from the reads, as well as repetitive and overlapping reads, because the findings in 1(b) are superior to 1(a). The HISAT2 tool was utilized to map the sequences to the human reference genome. This is a rapid and accurate method for mapping spliced reads to a genome. Next, in the absence of a reference transcriptome, we created transcriptomes using Stringtie that represented each neuroAIDS RNA-seq library. Additionally, we created a transcriptome database by merging redundant transcript structures from our sequences with our human RefSeq reference using the Stringtie - Merge tool. The transcripts of our newly constructed transcriptome were then annotated using GFFcompare (Trapnell et al., 2010) to show us how each transcript related to the human RefSeq reference, as shown in Table 1.

Table 1: GffCompare annotated transcripts (1st few lines of the output)

Seq name	Source	Feature	Start	End	Score	Strand	Frame	Attributes
chr 1	StringTie	transcript	1187404	1409	.	+	.	transcript_id "NR_046018.2"; gene_id "MSTRG.1"; gene_name "NR_046018.2"; xloc "XLOC_000001"; ref_gene_id "NR_046018.2"; cmp_ref "NR_046018.2"; class_code "="; tss_id "TSS1";
chr 1	StringTie	exon	1187404	1222407	.	+	.	transcript_id "NR_046018.2"; gene_id "MSTRG.1"; exon_number "1";
chr 1	StringTie	exon	1261603	1272401	.	+	.	transcript_id "NR_046018.2"; gene_id "MSTRG.1"; exon_number "2";
chr 1	StringTie	exon	1322201	1409	.	+	.	transcript_id "NR_046018.2"; gene_id "MSTRG.1"; exon_number "3";

Seq na me	So urc e	Fea tur e	St ar t	En d	Sc or e	St ra nd	Fr a m e	Attributes
chr 1	Stri ngT ie	tra nsc ript	14 41 7	29 35 3	.	+	.	transcript_id "MSTRG.3.2"; gene_id "MSTRG.3"; gene_name "NR_024540.1"; xloc "XLOC_000002"; cmp_ref "NR_024540.1"; class_code "s"; tss_id "TSS2";
chr 1	Stri ngT ie	exo n	14 41 7	14 82 9	.	+	.	transcript_id "MSTRG.3.2"; gene_id "MSTRG.3"; exon_number "1";
chr 1	Stri ngT ie	exo n	14 97 0	15 03 8	.	+	.	transcript_id "MSTRG.3.2"; gene_id "MSTRG.3"; exon_number "2";
chr 1	Stri ngT ie	exo n	15 79 6	16 76 5	.	+	.	transcript_id "MSTRG.3.2"; gene_id "MSTRG.3"; exon_number "3";
chr 1	Stri ngT ie	exo n	16 87 6	17 05 5	.	+	.	transcript_id "MSTRG.3.2"; gene_id "MSTRG.3"; exon_number "4";
chr 1	Stri ngT ie	exo n	17 91 5	18 06 1	.	+	.	transcript_id "MSTRG.3.2"; gene_id "MSTRG.3"; exon_number "5";

Next, we used Feature Counts tool which counted the reads aligning in exons of our GFF Compare generated transcriptome database.



Further, by using the DESeq2 tool we performed differential gene expression analysis using the read counts produced by Feature Counts (Fig. 2).

Figure 2 PCA analysis of sequences

Since PC1 and PC2 are the primary components that capture the greatest and second-most variation, respectively, they are constructed in the order that they cover the most variation. Our figure (Fig. 2) shows that PC1 has a variance of 53%. The points at the top, which correspond to SRR1852861 and SRR1852860, respectively, may be the source of the most variance in our data set. Data 71, which corresponds to SRR1852881, and data 68, which corresponds to SRR1852859, may be explained by our PC2's 25% variance. Additionally, to uncover genes with substantial fold changes that are also statistically significant, we utilized the volcano plot (Fig. 3) to plot the statistical significance (P value) versus the size of change (fold change). These genes may be the most important biologically. They are statistically significant. In this plot, the most upregulated genes are towards the right, the most downregulated genes are towards the left, and the most statistically significant genes are towards the top (Doyle, 2021). We have selected the 1st 10 genes from the DESeq2 results (NR_002834.1, NR_168399.1,NM 024654.5, NM_005105.5, XR_001737741.2,NR_120331.1,XR_001738448.1, NR_026967.1, NR_153413.2 and XR_945148.2) to be displayed in the volcano plot as expressed and significant genes. NR_002834.1,NR168399.1, NM 024654.5,NM 005105.5,XR_001738448.1, NR_026967.1, NR_153413.2 are the most upregulated genes, which are towards the right, and XR_001737741.2 and NR_120331.1 are the most downregulated genes, which are towards the left. From the genes selected, NR_168399.1 is the most significant gene since it is at the topmost of the list of genes.

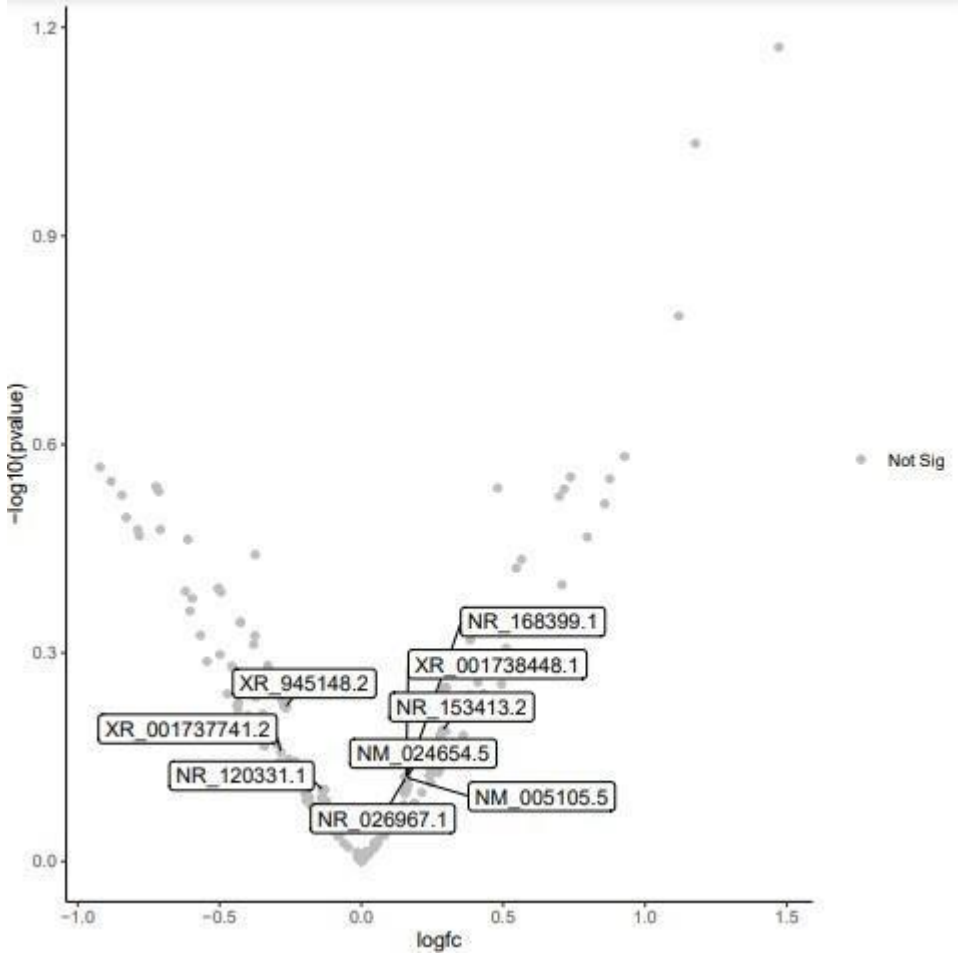


Figure 3 Volcano Plot

Structure based drug design

Further, for neuroAIDS, we go ahead with designing novel drugs from medicinal plants. The gene receptors for neuroAIDS corresponding to Fig. 7 are taken from NCBI for our work (Table 2).

Table 2: Genes with their NCBI Accession Number

Sl. No	Gene Receptors	NCBI Accession Number	Homologous Template
1.	AKT1	AAL55732.1	3QWQA
2	CCR5	AAB57793.1	4MBSA
3	MDM2	ABT17086.1	2LZGA

Genes Involved:

1. **AKT1:** Alpha serine/threonine-protein kinase
2. **CCR5:** C-C chemokine receptor type 5
3. **MDM2:** Mouse Double Minute 2 homolog

Homology modeling

The SWISS-MODEL server is used for the homology modeling of the aforementioned receptors. Figure 4 displays the receptor model and related Ramachandran plot results. Table 3 provides the modeling template that was employed.

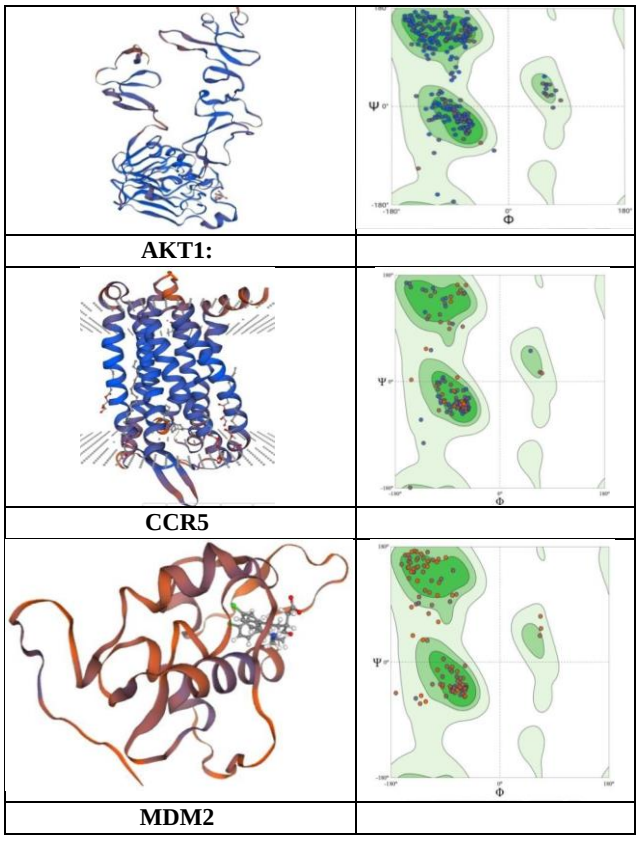


Figure 4 Swiss-model generated receptor models with their ramachandran plot

Ayurvedic Medicinal plants given in Table 5 are known to treat mental ailments and virus infections (Hegde, & Harini, 2014). The potency of their phytocompounds in treating neuroAIDS is studied here.

Table 3: Herbs used along with the active compounds (Phytocompounds)

Sr. No.	Medicinal Herbs	Active chemical Compounds
1.	Rheum species	1'-(phenylmethyl)
2.	*Veronia amygdali*	Caryophyllene oxide
3.	*Hypoxis hemerocallidea*	13Methyl6,7,8,9,11,12,14,15,16,17-decahydrocyclopenta[a]phenanthren-3-ol
		α-Myrcene
		α-d-Mannofuranoside, O-geranyl
4.	*Sutherlandia frutescens*	D-Pinitol
		4-Amino-3-hydroxybutyric acid
5.	*Hypericum perforatum L*	Benzanthrone
		Phloroglucinols hyperforin
6.	*Terminalia paniculata*	1,1-diethoxypropane
		Tert-Butyl hydrogen phthalate
		Cyclopropane carboxylic acid
7.	*Smilax corbularia* (Kunth)	gamma-Terpinene
		Sabinene
8.	*Astragalus membranaceus* (Bunge)	N-{5-[2-Chloro-5-(trifluoromethyl) phenyl] pyrazin-2-yl}-2,6-difluorobenzamide
9.	*Dittrichia viscosa* (L.) (Greuter)	Beta costic acid
		Costic acid;costus acid
		Costic acid methyl ester
10.	*Momordica balsamina*	5-p-Nitrobenzoyl gentisic acid
11.	*Calophyllum inophyllum*	Calanolide A
12.	*Syzygium claviflorum*	Ellagic acid
		Flavylium
		Kaempferol
13.	*Withania somnifera*(Ashwagandha)	Withaferin A

As per Lipinski's rule of five [ADME (adsorption, distribution, metabolism, and extraction)], we checked the drug likeliness of the phytocompounds listed in Table 5.

Molecular Docking:

Further docking is performed with the receptors in Table 4 and the phytocompounds listed in Table 5. Docking scores, interacting amino acids, and the number of interactions is noted in Tables7-9.

Table 4: Docking results of the AKT1 gene receptor

Sr. no	Receptor protein	Compound	Docking score (in – kcal/mol)	Interacting amino acids	No. of interactions
1.	AKT1	(1,1'-Biphenyl)-4,4'-	3724	THR-197	1

Sr. no	Receptor protein	Compound	Docking score (in – kcal/mol)	Interacting amino acids	No. of interactions
		diamine, 3-bromo-			
		1,1-Diethoxypropane	3248	THR-160	1
		4-Amino-3-hydroxybutyric acid	2440	ILE-290,TYR-229,ARG-206,LYS-419,GLU-228	5
		5-p-Nitrobenzoyl gentisic acid	4018	LYS-149, ARG-206, LYS-289	3
		13-Methyl-6,7,8,9,11,12,14,15,16,17-decahydrocyclopenta[a]phenanthren-3-ol	4212	0	0
		Benzanthrone	3650	TYR-326	1
		beta-Costic acid	3980	0	0
		-beta-d-Mannofuranoside, O-geranyl	4960	THR-195	1
		Buta-1,3-diene-1-sulfonyl chloride	3022	GLU-198, ILE-186, THR-195	3
		Calanolide A	4996	ARG-328, ASP- 398, GLY-395, ALA-50	4
		Caryophyllene oxide	4022	0	0
		Costic acid methyl ester	4114	LEU-78, GLN-59	2
		Costic acid; Costus acid	3946	ASN-279	3
		Cyclopropanecarboxylic acid	2156	ARG-15	1
		D-Pinitol	2922	GLU-298, GLU- 85	4
		Ellagic acid	3632	ARG-328, ASP- 325, ALA-50	5
		Flavylium	3798	0	0
		Gamma-Terpinene	3370	0	0
		Kaempferol	4044	LYS-149, TYR- 176, TYR-229	4
		Magnesium,1-(phenylmethyl)piperidine,br omide	3804	0	0
		Myrcene	3330	0	0
		Myristic acid	4276	SER-124, SER- 126, GLU-117	4
		N-{5-[2-Chloro-5-(trifluoromethyl)phenyl]pyr azin-2-yl}-	5362	GLU-198	1

Sr. no	Receptor protein	Compound	Docking score (in – kcal/mol)	Interacting amino acids	No. of interactions
		2,6- difluorobenzamide			
		Phloroglucinol	2468	THR-160, LYS- 297	3
		Sabinene	3276	0	0
		Tert-Butyl hydrogen phthalate	3768	LYS-276, TYR- 315, ASP-274, LEU-295	5
		Withaferin A	5626	TYR-229, GLU- 278, ASP-274	4

Table 5: Docking results of the CCR5 gene receptor

Sr. no	Receptor protein	Compound	DOCKING SCORE	Interacting amino acid	No. of interaction
2.	CCr5	(1,1'-Biphenyl)-4,4'-diamine, 3-bromo-	3388	SER-272	1
		1,1-Diethoxypropane	2738	GLN-280	1
		4-Amino-3-hydroxybutyric acid	2454	ASP-76, PHE-112	2
		5-p-Nitrobenzoyl gentisic acid	4156	ALA-92	1
		13-Methyl-6,7,8,9,11,12,14,15,16,17-decahydrocyclopenta[a]phenanthren-3-ol	4098	0	0
		Benzanthrone	3556	LYS-59	1
		beta-Costic acid	4056	0	0
		-beta-d-Mannofuranoside, O-geranyl	4682	GLN-280, TYR-37	2
		Buta-1,3-diene-1-sulfonyl chloride	2674	0	0
		Calanolide A	4954	ASP-66, THR-65, SER-63	4
		Caryophyllene oxide	3836	0	0
		Costic acid methyl ester	3920	0	0
		Costic acid; Costus acid	3902	GLU-302	1
		Cyclopropane carboxylic acid	2328	GLY-111, PHE-112	2
		D-Pinitol	2734	LYS-303, GLU-302	2
		Ellagic acid	3568	GLU-302, LYS-59, ASP-66	4
		Flavylium	3758	0	0
		Gamma-Terpinene	3136	0	0
		Kaempferol	3768	ARG-232, GLU-302	2
		Magnesium,1- (phenylmethyl)piperidine, bromide	3466	0	0
		Myrcene	3250	0	0
		Myristic acid	4250	GLN-280, THR-284	2
		N-{5-[2-Chloro-5-(trifluoromethyl) phenyl]	5266	TYR-187	1

Sr. no	Receptor protein	Compound	DOCKING SCORE	Interacting amino acid	No. of interaction
		pyrazin-2-yl}-2,6- difluorobenzamide			
		Phloroglucinol	2376	0	0
		Sabinene	3030	0	0
		Tert-Butyl hydrogen phthalate	3588	ASP-66, LYS-59	4
		Withaferin A	5444	TYR-37	1

Table 6: Docking results of MDM2 gene receptor

Sr. no	Receptor protein	Compound	DOCKIN G SCORE	Interacting amino acid	No. of interaction
3.	MDM2	(1,1'-Biphenyl)-4,4'-diamine, 3- bromo-	3028	THR-26	1
		1,1-Diethoxypropane	2828	THR-26	1
		4-Amino-3-hydroxybutyric acid	1974	GLN-24, ASN- 111, THR-26, TYR-104	4
		5-p-Nitrobenzoyl gentisic acid	3812	GLU-23, GLU-25, THR-26, ASN-111, GLN- 117, SER-116	11
		13-Methyl- 6,7,8,9,11,12,14,15,16,17-decahydrocyclopenta[a] phenanthren-3-ol	3784	ARG-29	1
		Benzanthrone	3316	0	0
		beta-Costic acid	3524	GLN-112	1
		-beta-d-Mannofuranoside, O- geranyl	4268	GLN-113, GLU- 114	4
		Buta-1,3-diene-1-sulfonyl chloride	2478	THR-26, ASN- 111	2
		Calanolide A	4632	GLU-114, GLN- 113	3
		Caryophyllene oxide	3366	GLN-24	1
		Costic acid methyl ester	3696	THR-26	1
		Costic acid; Costus acid	3654	0	0
		Cyclopropane carboxylic acid	1750	0	0
		D-Pinitol	2470	GLN-113, ASN- 111, THR-26	3
		Ellagic acid	3458	GLN-24, GLU- 25, THR-26, GLN-113, GLU- 114, GLN-112	7
		Flavylium	3552	0	0
		Gamma-Terpinene	2818	0	0
		Kaempferol	3578	THR-26	1
		Magnesium,1- (phenylmethyl)piperidine, bromide	3072	0	0
		Myrcene	2874	0	0
		Myristic acid	3794	0	0
		N-{5-[2-Chloro-5-(trifluoromethyl) phenyl] pyrazin-2-yl}-2,6- difluorobenzamide	5246	GLN-112, ASP- 117	2
		Phloroglucinol	1964	0	0
		Sabinene	2844	0	0
		Tert-Butyl hydrogen phthalate	3338	THR-120, ASP- 117, GLY-119, SER-118	6
		Withaferin A	5086	GLN-113, GLU- 114	2

It is seen that the AKT1 receptor docks with good interactions with 4-Amino-3-hydroxybutyric acid

with a docking score of -2440 kcal/mol, 5-p-Nitrobenzoyl gentisic acid with a docking score of – 4018 kcal/mol, Buta-1,3-diene-1-sulfonyl chloride with a docking score of -3022 kcal/mol, Calanolide A with a docking score of -4996 kcal/mol, D-Pinitol with a docking score of – 2922 kcal/mol, Ellagic acid with a docking score of -3632 kcal/mol, Kaempferol with a docking score of -4044 kcal/mol, Myristic acid, with a docking score of -4276 kcal/mol, Tert-Butyl hydrogen phthalate with a docking score of -3768kcal/mol and Withaferin A with a docking score of -5626 kcal/mol.

Again, it is seen that CCr5 receptor docks with a good interaction with Calanolide A with a docking score of -4954 kcal/mol, Ellagic acid with a docking score of -3568kcal/mol and Tert- Butyl hydrogen phthalate with a docking score of -3588 kcal/mol.

Further, it is seen that MDM2 receptor docks with a good interactions with 4-Amino-3- hydroxybutyric acid with a docking score of -1974 kcal/mol, 5-p-Nitrobenzoyl gentisic acid with a docking score of -3812 kcal/mol, beta-d-Mannofuranoside, O-geranyl with a docking score of – 4268 kcal/mol, Calanolide A with a docking score of -4632 kcal/mol, Ellagic acid with a docking score of -3458 and Tert-Butyl hydrogen phthalate with a docking score of -3338kcal/mol.

4. Conclusion

The study used MultiQC and DESeq tools to analyze gene expression dynamics in a sample. Trimmomatic was used for quality control, removing overlapping and low-quality sequences to ensure high-quality data for subsequent analyses. DESeq analysis identified active genes, such as NR_002834.1, NR_168399.1, NM_024654.5, NM_005105.5, XR_001738448.1, NR_026967.1, and NR_153413.2, which likely play significant roles in the biological processes under investigation. Decreased activity genes, such as XR_001737741.2 and NR_120331.1, were identified, suggesting potential regulatory mechanisms or involvement in disease pathways. LOC107985115 variant 4 was identified as a key gene, possibly playing a central role in the biological processes being studied. The integrated analysis provides a comprehensive understanding of gene expression patterns, revealing potential molecular mechanisms underlying biological processes and paving the way for further research.

From the docking studies, it is seen that Calanolide A, Ellagic acid and Tert-Butyl hydrogen phthalate dock with all three receptors. Hence, it can be concluded that Calanolide A, Ellagic acid and Tert-Butyl hydrogen phthalate act as ligands for AKT1, CCR5, and MDM2 receptors. Further, *in-vitro* and *in-vivo* studies can be done on these phytocompounds to confirm their efficacy as a drug for neuroAIDS.

5. Reference

Alkhatib, G. (2009). The biology of CCR5 and CXCR4. *Current Opinion in HIV and AIDS*, *4*(2), 96-103. https://doi.org/10.1097/COH.0b013e328324bbec

Ances, B. (2008). Emerging views of NeuroAIDS. *The Lancet Neurology*, *7*(7), Article 580. https://doi.org/10.1016/S1474-4422(08)70134-0

Andrews, S. (2000). *FastQC A Quality Control tool for High Throughput Sequence Data*. Retrieved from http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Batut, B., Hiltemann, S., Bagnacani, A., Baker, D., Bhardwaj, V., Blank, C., ... & Grüning, B. (2018). Community-driven data analysis training for biology. *Cell systems*, *6*(6), 752-758. https://doi.org/10.1016/j.cels.2018.05.012

Blankenberg, D., & Hillman-Jackson, J. (2014). Analysis of next-generation sequencing data using Galaxy. *Stem Cell Transcriptional Networks: Methods and Protocols*, 21-43. https://doi.org/10.1007/978-1-4939-0512-6_2

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114-2120. https://doi.org/10.1093/bioinformatics/btu170

Bradshaw, R. A., & Stahl, P. D. (2015). *Encyclopedia of cell biology*. Academic Press.

Doyle M. (2021). *Visualization of RNA-Seq results with Volcano Plot (Galaxy Training Materials)*. Retrieved

from https://training.galaxyproject.org/training-material/topics/transcriptomics/tutorials/rna-seq-viz-with-volcanoplot/tutorial.html

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, *32*(19), 3047-3048. https://doi.org/10.1093/bioinformatics/btw354

Felger, J. C., & Treadway, M. T. (2017). Inflammation effects on motivation and motor activity: role of dopamine. *Neuropsychopharmacology*, *42*(1), 216-241. https://doi.org/10.1038/npp.2016.143

Fernandez-Cruz, A. L., & Fellows, L. K. (2017). The electrophysiology of neuroHIV: A systematic review of EEG and MEG studies in people with HIV infection since the advent of highly-active antiretroviral therapy. *Clinical Neurophysiology*, *128*(6), 965-976. https://doi.org/10.1016/j.clinph.2017.03.035

Hegde, P. L. & Harini, A. (2014), A text book of Dravyaguna Vijnana, Chaukhambha Publications.

Kaul, M., Zheng, J., Okamoto, S., Gendelman, H. E., & Lipton, S. A. (2005). HIV-1 infection and AIDS: consequences for the central nervous system. *Cell Death & Differentiation*, *12*(1), 878-892. https://doi.org/10.1038/sj.cdd.4401623

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nature methods*, *12*(4), 357-360. https://doi.org/10.1038/nmeth.3317

Kovaka, S., Zimin, A. V., Pertea, G. M., Razaghi, R., Salzberg, S. L., & Pertea, M. (2019). Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome biology*, *20*, 1-13. https://doi.org/10.1186/s13059-019-1910-1

Kranick, S. M., & Nath, A. (2012). Neurologic complications of HIV-1 infection and its treatment in the era of antiretroviral therapy. *CONTINUUM: Lifelong Learning in Neurology*, *18*(6), 1319-1337.

Li, W., Freudenberg, J., Suh, Y. J., & Yang, Y. (2014). Using volcano plots and regularized-chi statistics in genetic association studies. *Computational biology and chemistry*, *48*, 77-83. https://doi.org/10.1016/j.compbiolchem.2013.02.003

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*(7), 923-930. https://doi.org/10.1093/bioinformatics/btt656

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, *15*, 1-21. https://doi.org/10.1186/s13059-014-0550-8

Lowe, R., Shirley, N., Bleackley, M., Dolan, S., & Shafee, T. (2017). Transcriptomics technologies. *PLoS computational biology*, *13*(5), e1005457. https://doi.org/10.1371/journal.pcbi.1005457

McArthur, J. C. (1987). Neurologic manifestations of AIDS. *Medicine*, *66*(6), 407-437. https://doi.org/10.1097/00005792-198711000-00001

McCombe, J. A., Noorbakhsh, F., Buchholz, C., Trew, M., & Power, C. (2009). NeuroAIDS: a watershed for mental health and nervous system disorders. *Journal of Psychiatry and Neuroscience*, *34*(2), 83-85.

Minagar, A., Commins, D., Alexander, J. S., Hoque, R., Chiappelli, F., Singer, E. J., ... & Shapshak, P. (2008). NeuroAIDS: characteristics and diagnosis of the neurological complications of AIDS. *Molecular diagnosis & therapy*, *12*, 25-43. https://doi.org/10.1007/BF03256266

Molinspiration Cheminformatics. (n.d.). Retrieved from https://www.molinspiration.com

Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., & Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature protocols*, *11*(9), 1650-1667. https://doi.org/10.1038/nprot.2016.095

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., & Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature biotechnology*, *33*(3), 290-295. https://doi.org/10.1038/nbt.3122

Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., & Wolfson, H. J. (2005). PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic acids research*, *33*(suppl_2), W363-W367. https://doi.org/10.1093/nar/gki481

Shapshak, P., Kangueane, P., Fujimura, R. K., Commins, D., Chiappelli, F., Singer, E., ... & Sinnott, J. T. (2011). Editorial neuroAIDS review. *Aids*, *25*(2), 123-141. https://doi.org/10.1097/QAD.0b013e328340fd42

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., Van Baren, M. J., ... & Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature biotechnology*, *28*(5), 511-515. https://doi.org/10.1038/nbt.1621

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., ... & Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic acids research*, *46*(W1), W296-W303. https://doi.org/10.1093/nar/gky427

Wilen, C. B., Tilton, J. C., & Doms, R. W. (2012). HIV: cell binding and entry. *Cold Spring Harbor perspectives in medicine*, *2*(8), a006866. https://doi.org/10.1101/cshperspect.a006866