

OPEN

# Novel pathways of HIV latency reactivation revealed by integrated analysis of transcriptome and target profile of bryostatin

Bing-xiang Li<sup>1,6</sup>, Han Zhang<sup>1,6</sup>, Yubin Liu<sup>3,6</sup>, Ya Li<sup>4,6</sup>, Jun-juan Zheng<sup>5</sup>, Wen-Xing Li<sup>5</sup>, Kai Feng<sup>1</sup>, Ming Sun<sup>1\*</sup> & Shao-Xing Dai<sup>2\*</sup>

The reactivation of HIV latency cell will be necessary to curing HIV infection. Although many latency-reversal agents (LRAs) have proven effective to reactivate the latency cell, there is a lack of any systematic analysis of the molecular targets of these LRAs and related pathways in the context of transcriptome. In this study, we performed an integrated analysis of the target profile of bryostatin and transcriptome of the reactivated CD4<sup>+</sup> T cells after exposing to bryostatin. The result showed a distinct gene expression profile between latency cells and bryostatin reactivated cells. We found bryostatin can target multiple types of protein other than only protein kinase C. Functional network analysis of the target profile and differential expressed genes suggested that bryostatin may activate a few novel pathways such as pyrimidine metabolism, purine metabolism and p53 signaling pathway, besides commonly known pathways DNA replication, cell cycle and so on. The results suggest that bryostatin may reactivate the HIV-latent cells through up-regulation of pyrimidine and purine metabolism or through starting the cell-cycle arrest and apoptosis induced by up-regulation of p53 signaling pathway. Our study provides some novel insights into the role of bryostatin and its affected pathways in controlling HIV latency and reactivation.

Human immunodeficiency virus type 1 (hereafter referred to as “HIV”) is the causative agent of acquired immune deficiency syndrome (AIDS). In the host cells, HIV infection causes to produce viral RNAs and proteins, results in profound functional and morphological changes, and alerts the progression of the host cell cycle. Meanwhile, HIV genomes can integrate to the host cellular DNA as a persistent viral reservoir and harboring replication-competent proviruses with a dormancy state<sup>1</sup>.

Through the development of specific inhibitors of HIV replication, significant progress has been made in the treatment of HIV infection. The success of anti-retroviral (ARV) therapy reduces HIV plasma viremia to an under-detectable level<sup>2</sup>. However, HIV is not eradicated and AIDS patients must remain on ARV drugs for life. A major obstacle to cure AIDS is the HIV latency that is a state of reversibly nonproductive infection in the resting CD4 T cells and possibly other cells<sup>3</sup>. The long-lived reservoir is resistant to ARV therapy and to clearance by the host immune system. Once treatment is stopped, HIV originating from the reservoir lead to re-seed the infection<sup>4</sup>.

The molecular mechanisms of HIV latency are complex and numerous studies have been performed to explain this phenomenon<sup>3,5,6</sup>. In general, a combination of trans-effects and cis-effects governs the activity of HIV promoter transcription making the reservoir the best-established barrier to HIV eradication. Mostly cis-effects

<sup>1</sup>Institute of Medical Biology, Peking Union Medical College and Chinese Academy of Medical Sciences, Kunming, China. <sup>2</sup>Yunnan Key Laboratory of Primate Biomedicine Research, Institute of Primate Translational Medicine, Kunming University of Science and Technology, Kunming, 650500, Yunnan, China. <sup>3</sup>Department of Infectious Diseases, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing, China. <sup>4</sup>Yunnan Key Laboratory of Laboratory Medicine, Yunnan Institute of Laboratory Diagnosis, Department of Clinical Laboratory, The First Affiliated Hospital of Kunming Medical University, Kunming, 650032, Yunnan, China. <sup>5</sup>State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, 650223, Yunnan, China. <sup>6</sup>These authors contributed equally: Bing-xiang Li, Han Zhang, Yubin Liu and Ya Li. \*email: [sunming@imbcams.com.cn](mailto:sunming@imbcams.com.cn); [daishx08@gmail.com](mailto:daishx08@gmail.com)

reflect the variety of chromatin environments at different sites of integration genome and trans-effects reflect the trans-acting transcription factors within the host cells<sup>5</sup>. For example, heterochromatin of the host genome after HIV infection impairs the gene expression by impeding transcription factors access to the underlying DNA<sup>7</sup>. As the fundamental structural unit of chromatin, nucleosomes are regulated by posttranslational modifications of the histone tails which include phosphorylation, methylation, acetylation, ubiquitylation and poly ADP-ribosylation<sup>8</sup>. Specific combinations of these repressive histone marks have been proposed to govern the chromatin assembly and HIV gene expression<sup>9</sup>. Similarly, the strong repressive state of the HIV latency is correlated with methylation of CpG islands within the HIV promoter and DNA demethylation is associated with re-activation of latent HIV<sup>10</sup>. Another mechanism that can cause HIV transcriptional silence is the transcriptional interference of the HIV promoters and cellular promoters in the integration genome<sup>11</sup>. Depending on the orientation of the HIV provirus relative to the host cells genes, different forms of transcriptional interference occur<sup>12</sup>.

It has become clear that reactivation of HIV latency cells in AIDS patients on ARV therapy will be necessary to curing HIV infection. This is the strategy named “shock and kill” to facilitate a cure for HIV. This strategy administers latency reactivation agents to patients while on ARV therapy in order to induce pro-viral transcription in latently infected cells<sup>13</sup>. Several classes of compounds have been discovered to “Shock” HIV out of latency including; protein kinase C (PKC) agonists<sup>14</sup>, histone deacetylase inhibitors (HDACis)<sup>15</sup>, histone methyltransferase inhibitors (HMTIs)<sup>16</sup>, cytokines, bromodomain inhibitors<sup>17</sup> and so on.

The PKC agonists is the only one latency-reversal agents (LRAs) class to demonstrate reproducible results across different HIV latency models<sup>18</sup>. Generally, the PKC agonists can be divided into three major chemical families: phorbol esters, including prostratin, 12-deoxyphorbol 13-phenylacetate (DPP) and phorbol 12-myristate 13-acetate (PMA); diterpenes, which include compounds of ingenol; and macrocyclic lactones including bryostatin and its analogs<sup>19</sup>. Among these PKC agonists, bryostatin-1 (hereafter referred to as “bryostatin”) has shown significant potency to revert HIV latency in *ex vivo* experiment using patient cells, compared with other LRAs<sup>20</sup>.

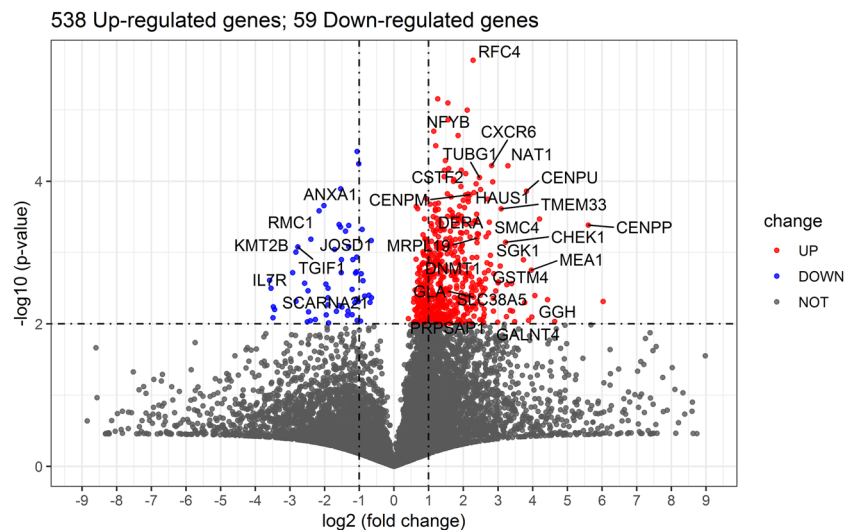
Although these compounds above have proven effective to reactivate the latency cell, the molecular mechanisms underlying the effects are not entirely clear. The activities of these compounds cannot be attributed wholly to a single target and may involve several molecular targets and pathways. Unfortunately, there is no systematic analysis of molecular targets of these LRAs and their related pathway. In this study, we performed an integrated analysis of the target profile of bryostatin and transcriptome of the reactivated CD4<sup>+</sup> T cells after exposing to bryostatin. The result showed a distinct gene expression profile between the latency cells and bryostatin reactivated cells. Extensive changes of gene expression occurred in the CD4<sup>+</sup> T cells treated with bryostatin. Furthermore, we found bryostatin can target multiple types of protein other than only PKC. Functional network analysis of the target profile and differential expressed genes (DEGs) suggested that bryostatin may activate a few novel pathways such as pyrimidine metabolism and purine metabolism and p53 signaling pathway, in addition to the commonly known pathways DNA replication, cell cycle, nucleotide excision repair and mismatch repair. These results provide mechanistic insights into the role of bryostatin and its affected pathways in controlling HIV latency and reactivation.

## Results

**Extensive changes of gene expression in CD4<sup>+</sup> T cell exposed to bryostatin.** Compared with the unstimulated CD4<sup>+</sup> T cell, we identified 597 DEGs (P value < 0.01, Table S1) in the bryostatin stimulated CD4<sup>+</sup> T cell. In the DEGs, there are 538 up-regulated and 59 down-regulated genes (Fig. 1). It indicates that extensive changes of gene expression in CD4<sup>+</sup> T cells after exposing to bryostatin. We listed the top 30 most significantly deregulated genes, most of which were up-regulated (Table 1 and labeled in Fig. 1). The expression of genes *CENPP*, *GGH*, *MEA1*, *CENPU*, *SGK1*, *GALNT4*, and *SLC38A5* was upregulated more than 10 folds. In contrast, the expression of genes *KMT2B*, *IL7R* was decreased by over four-fifths. The most obvious DEGs for up-regulation and down-regulation were *CENPP* (FC = 48.57) and *IL7R* (FC = 0.132), respectively.

**The bryostatin treated and unstimulated CD4<sup>+</sup> T cells showed a distinct gene expression profile.** To further overview the expression profile of CD4<sup>+</sup> T cells under two conditions, hierarchical clustering analysis with a heatmap was performed (Fig. 2). The hierarchical clustering was generated using 464 DEGs (P value < 0.01 and  $|\log_2FC| > 1$ ). In the hierarchical clustering result, 18 samples were divided two distinct categories clearly. The bryostatin treated samples showed a distinct gene expression profile from that of the unstimulated CD4<sup>+</sup> T cells, except the sample “TM3\_ unstimulated”. Most of these DEGs are up-regulated in the bryostatin treated CD4<sup>+</sup> T cells.

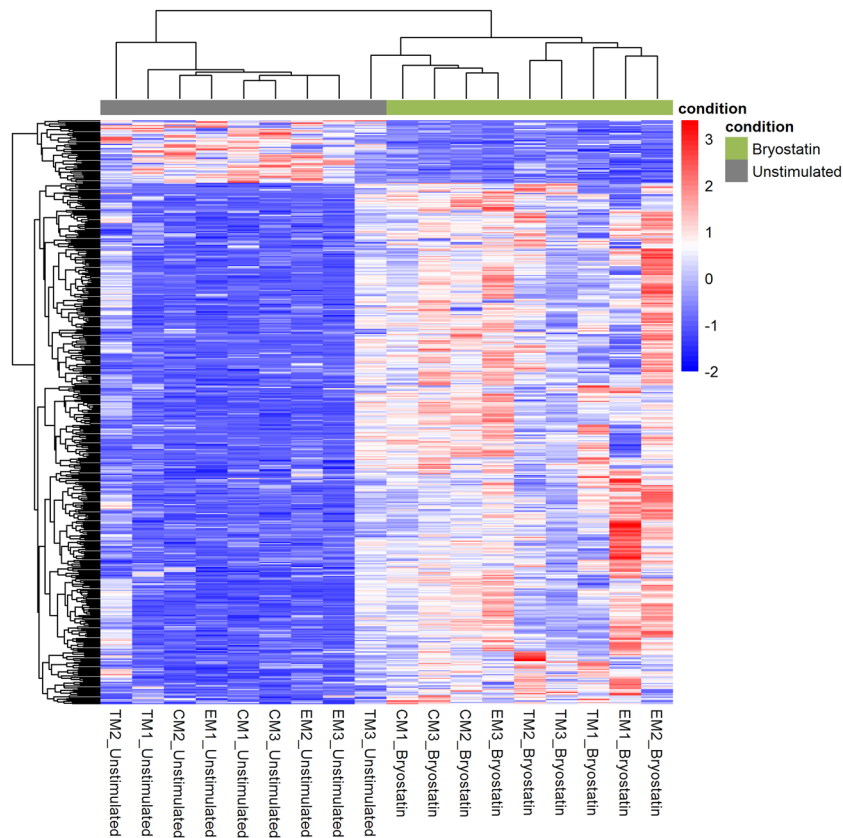
**Significantly changed pathways in CD4<sup>+</sup> cells from latency to reactivation state.** As mentioned above, after treated with bryostatin, CD4<sup>+</sup> cells were activated and extensive changes of gene expression occurred. A total of 538 and 59 genes were up-regulated and down-regulated, respectively. These significantly DEGs were enriched in 20 pathways revealed by KEGG pathway enrichment analysis and false discovery rate (FDR) adjustment (Fig. 3A). The top changed pathways were DNA replication, pyrimidine metabolism, drug metabolism - other enzymes, cell cycle, nucleotide excision repair, purine metabolism, mismatch repair, carbon metabolism, p53 signaling pathway, and biosynthesis of amino acids. In these pathways, 62 genes were significantly differential expressed (fold change > 2) (Fig. 3B). These genes can perform functions in multiple pathways. For example, *POLD2* gene is involved in five pathways such as purine metabolism, pyrimidine metabolism, DNA replication, nucleotide excision repair and mismatch repair. Another example is *PCNA* gene, which is involved in four pathways as showed in the Fig. 3B. These pathways and involved DEGs provide clear clues for the molecular mechanism of HIV latency reactivation.



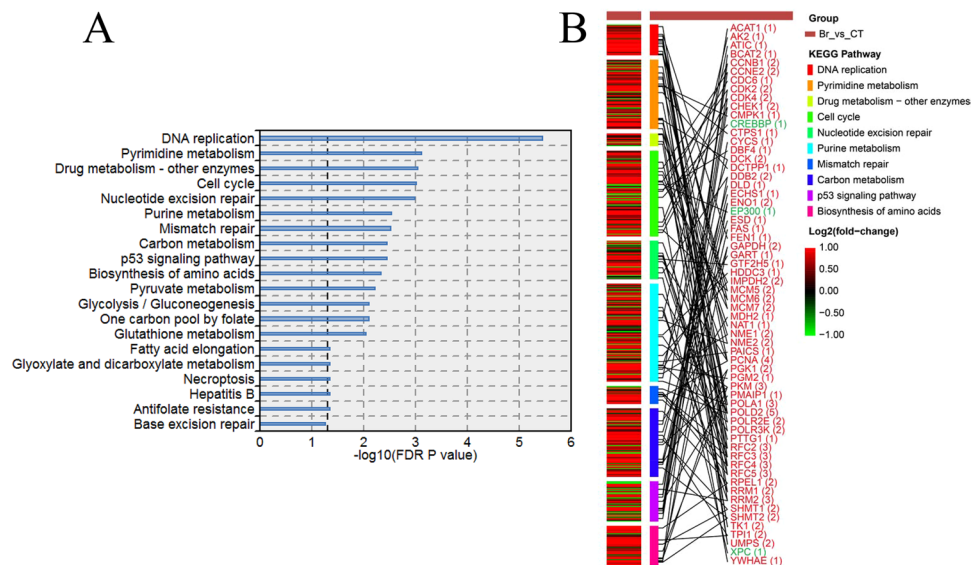
**Figure 1.** The differential expressed genes in CD4<sup>+</sup> T cell treated with bryostatin. The x-axis is log<sub>2</sub> ratio of gene expression levels between the bryostatin stimulated and unstimulated CD4<sup>+</sup> T cells; the y-axis is P value based on  $-\log_{10}$ . The red and blue dots represent the up-regulated and down-regulated gene (P value < 0.01), respectively; the most significant DEGs (P value < 0.05, AUC  $\geq$  0.95 and  $|\log_2FC| > 2$ ) were labeled with gene symbol.

Gene	Fold change	Log <sub>2</sub> (FC)	AUC	P value	Change trend
<i>CENPP</i>	48.5768	5.6022	1.0000	4.12E-04	UP
<i>GGH</i>	15.7602	3.9782	0.9753	8.00E-03	UP
<i>MEA1</i>	15.5598	3.9597	0.9877	1.76E-03	UP
<i>CENPU</i>	14.1561	3.8234	0.9877	1.37E-04	UP
<i>SGK1</i>	13.2958	3.7329	0.9753	1.26E-03	UP
<i>GALNT4</i>	11.1719	3.4818	0.9753	9.35E-03	UP
<i>SLC38A5</i>	10.2226	3.3537	0.9753	6.48E-03	UP
<i>NAT1</i>	9.7619	3.2872	0.9877	6.06E-05	UP
<i>GSTM4</i>	9.5371	3.2536	0.9506	2.82E-03	UP
<i>CHEK1</i>	9.3069	3.2183	0.9506	7.16E-04	UP
<i>TMEM33</i>	8.5065	3.0886	0.9877	2.43E-04	UP
<i>CXCR6</i>	7.0550	2.8186	0.9753	6.03E-05	UP
<i>DNMT1</i>	6.8638	2.7790	0.9630	1.09E-03	UP
<i>SMC4</i>	6.7204	2.7485	0.9753	3.76E-04	UP
<i>TUBG1</i>	5.5322	2.4678	0.9506	8.85E-05	UP
<i>MRPL19</i>	5.2580	2.3945	0.9506	6.28E-04	UP
<i>CSTF2</i>	5.1697	2.3701	0.9630	1.08E-04	UP
<i>HAUS1</i>	4.9499	2.3074	0.9630	1.44E-04	UP
<i>PRPSAP1</i>	4.9473	2.3067	0.9506	7.81E-03	UP
<i>GLA</i>	4.9143	2.2970	0.9506	4.22E-03	UP
<i>RFC4</i>	4.8528	2.2788	0.9753	2.00E-06	UP
<i>CENPM</i>	4.5732	2.1932	0.9630	1.55E-04	UP
<i>NPYB</i>	4.3205	2.1112	0.9877	1.01E-05	UP
<i>DERA</i>	4.0983	2.0350	0.9753	2.33E-04	UP
<i>ANXA1</i>	0.2462	-2.0224	1.0000	2.19E-04	DOWN
<i>RMC1</i>	0.2233	-2.1627	0.9753	2.60E-04	DOWN
<i>JOSD1</i>	0.1898	-2.3978	0.9506	6.49E-04	DOWN
<i>SCARNA21</i>	0.1790	-2.4816	0.9506	3.44E-03	DOWN
<i>TGIF1</i>	0.1461	-2.7745	0.9630	8.33E-04	DOWN
<i>KMT2B</i>	0.1403	-2.8331	0.9877	9.87E-04	DOWN
<i>IL7R</i>	0.1319	-2.9230	0.9877	1.91E-03	DOWN

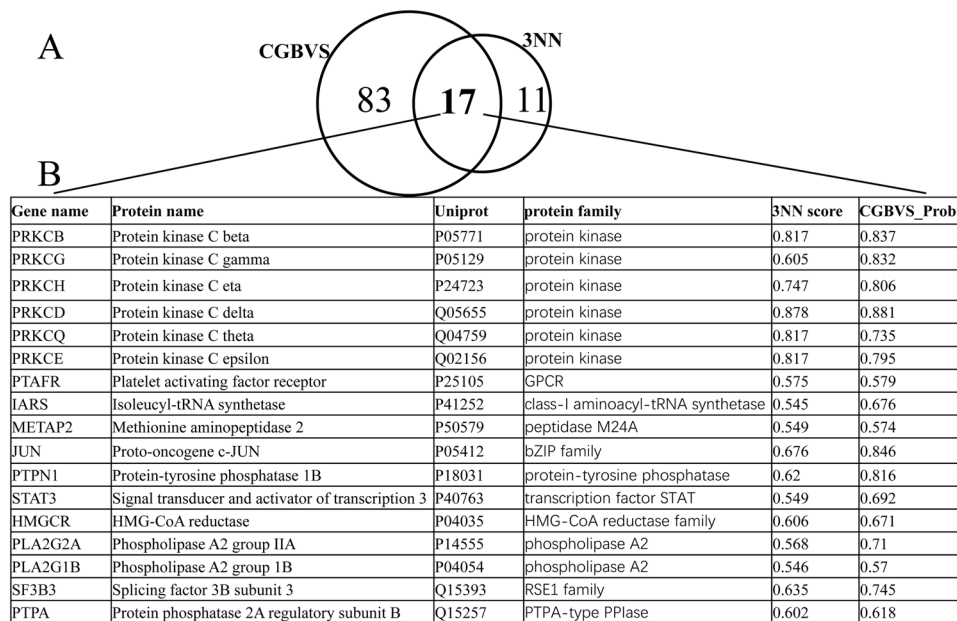
**Table 1.** The information about the most significantly deregulated genes ( $|\log_2 FC| > 2$  and AUC > 0.95).



**Figure 2.** A distinct gene expression profile between the bryostatin stimulated and unstimulated CD4<sup>+</sup> T cells. The heatmap of 464 significant DEGs (P value < 0.01 and |log<sub>2</sub>FC| > 1) shows the hierarchical clustering of the 18 samples. (**Bryostatin:** bryostatin treated cell; **Unstimulated:** unstimulated CD4<sup>+</sup> T cells). The color and shade change correspond to the expression value of gene after computing logarithms (log<sub>2</sub>) and normalization.



**Figure 3.** KEGG pathways significantly enriched by DEGs and the network of the top 10 enriched pathways and DEGs. (A) The KEGG pathways enriched by DEGs. The line width indicates the enrichment percentage. The dotted line in the box indicates the significance threshold (FDR P value = 0.05). P values were measured by a hypergeometric test. (B) The selected top 10 pathways and DEGs for the explanations the treatment effect of bryostatin. The network showed the relationship between the affected pathways and DEGs (|log<sub>2</sub>FC| > 2).

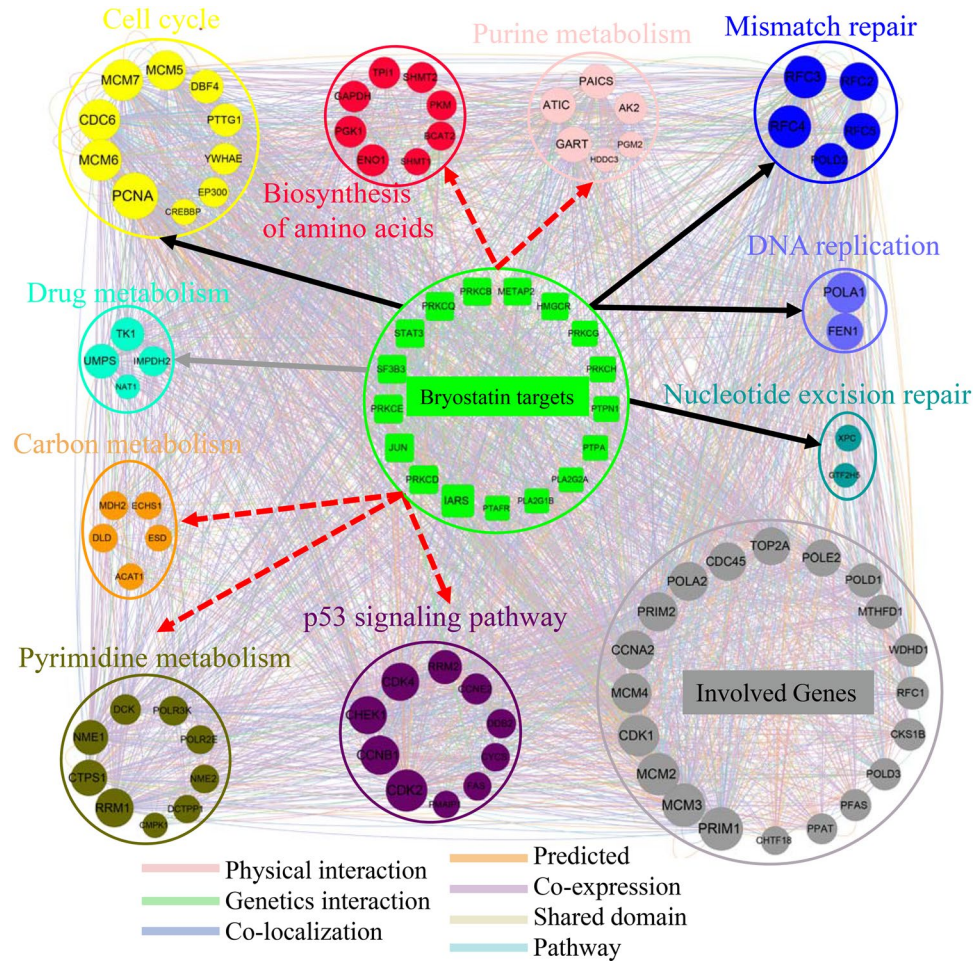


**Figure 4.** Venn diagram of the predicted targets for bryostatin by the two different methods and the consistent targets. **(A)** Venn diagram of the predicted targets by CGBVS and 3NN. **(B)** Details for the 17 consistent targets predicted by the two methods.

**The 17 putative targets of bryostatin predicted by two different methods.** In order to further understand the action mechanism of bryostatin, we predicted the putative targets for bryostatin. Target prediction is still a challenging task so far and plays a critical role in the drug development<sup>21</sup>. It is useful to explore ligand-target interactions and their related bio-chemical mechanisms. We used the two methods CGBVS and 3NN to predict the putative targets of bryostatin. Chemical genomics-based virtual screening (CGBVS) is a proteome-wide method which predicts compound and protein interactions (CPIs) only based on the character description of protein sequence. In contrast, 3NN method is based on ligand similarity. A total of 100 and 28 targets of bryostatin were predicted by CGBVS and 3NN, respectively (Fig. 4A). Among these predicted targets, 17 overlap targets were considered as consistent targets of bryostatin. The 17 overlap targets were *PRKCB*, *PRKCG*, *PRKCH*, *PRKCD*, *PRKCQ*, *PRKCE*, *PTAFR*, *IARS*, *METAP2*, *JUN*, *PTPN1*, *STAT3*, *HMGCR*, *PLA2G2A*, *PLA2G1B*, *SF3B3*, *PTPA* (Fig. 4B). They mainly belong to the protein families of protein kinase, phospholipase A2, GPCR, peptidase M24A and so on. Among these targets, protein kinase C family is the primary known targets of bryostatin. The other protein families are defined as putative targets of bryostatin based on the result of the two methods of CGBVS and 3NN methods. Bryostatin may reactivate the HIV latent CD4<sup>+</sup> cells through multiple targets and their affected pathways.

**Novel pathways of HIV latency reactivation revealed by functional network between bryostatin targets and DEGs.** KEGG enrichment analysis revealed the top 10 changed pathways and 62 genes in these pathways significantly differential expressed (fold change >2) after exposing to bryostatin. Meanwhile, we predicted 17 consistent targets for bryostatin. To determine whether or how these targets of bryostatin affect the top 10 changed pathways and DEGs, we constructed a functional network between these targets and DEGs using GeneMANIA (Fig. 5). As Fig. 5 showed, the functional network revealed intense interaction between bryostatin targets and DEGs. Based on the KEGG database, the DEGs were labeled by different pathways. We can see clearly that bryostatin reactivate the HIV latent CD4<sup>+</sup> cells through the effects of multiple targets. These targets affect several well-known pathways involved in the HIV latent reactivation such as DNA replication, cell cycle, nucleotide excision repair, and mismatch repair. Beside those well-known pathways, we also observed some novel pathways are closely linked to HIV latency reactivation (Fig. 5). These novel pathways are p53 signaling pathway and metabolic pathways including pyrimidine metabolism, purine metabolism, carbon metabolism, and biosynthesis of amino acids. They were not reported to be associated with latent HIV activation based on the past studies. The bryostatin targets may affect these metabolic pathways and p53 signaling, and thus play a role in the reactivation of HIV latent CD4<sup>+</sup> cells.

There are more than 30 genes in pyrimidine metabolism pathway highly expressed in the CD4<sup>+</sup> cells after exposing to bryostatin (Fig. 6). In contrast, only a few genes were down-regulated. From the upstream genes (*UMPS*, *DHODH*, *DCK*, *NME1*, *RRM1*, *PNP*, etc.) to downstream genes (*POLAI*, *POLR2E*, etc.) of this pathway, mostly of genes were highly active. Interestingly, we have found that the half parts of DNA synthesis and RNA transcription are highly active. In contrast, the other parts related to the catabolism and amino acid synthesis are under extreme inhibition and haven't changed much in gene expression. Similarly, more than 50 genes in purine metabolism pathway were highly expressed in the CD4<sup>+</sup> cells after exposing to bryostatin (Fig. 7). The pattern of gene expression was also similar to that of the pyrimidine metabolism pathway. This differential gene expression



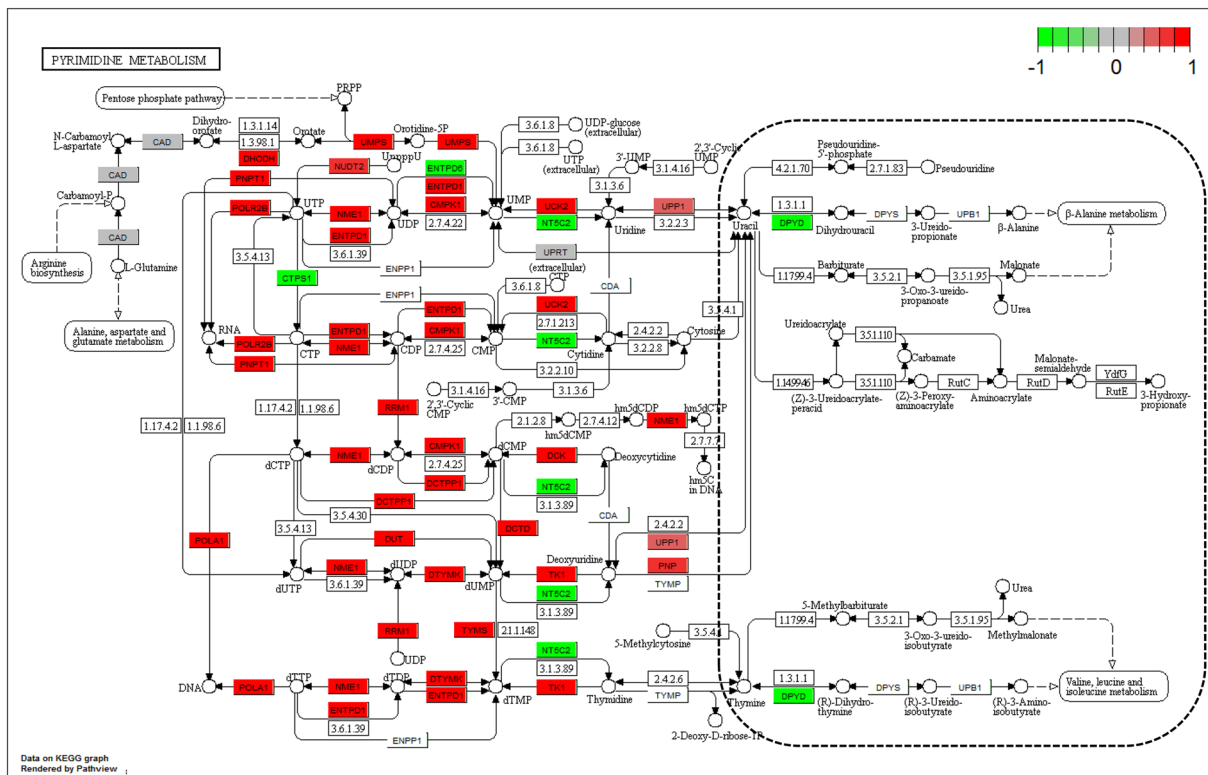
**Figure 5.** Functional network analysis between the consistent targets of bryostatin and the most DEGs. The network shows the functional interaction between the consistent targets (green rectangle) and most DEGs that are circled and labeled by different pathways. The gray nodes represent the result genes involved in the functional interaction between targets (green rectangle) and DEGs. The black arrows represent the known pathways involved in HIV latent reactivation. The dotted arrows represent the novel pathway revealed by this study. Each colored line represents a different interaction, and the color line width indicates the confidence of the interactions.

may drive the metabolic flux toward to the parts of DNA synthesis and RNA transcription, other than the parts related to the catabolism and amino acid synthesis.

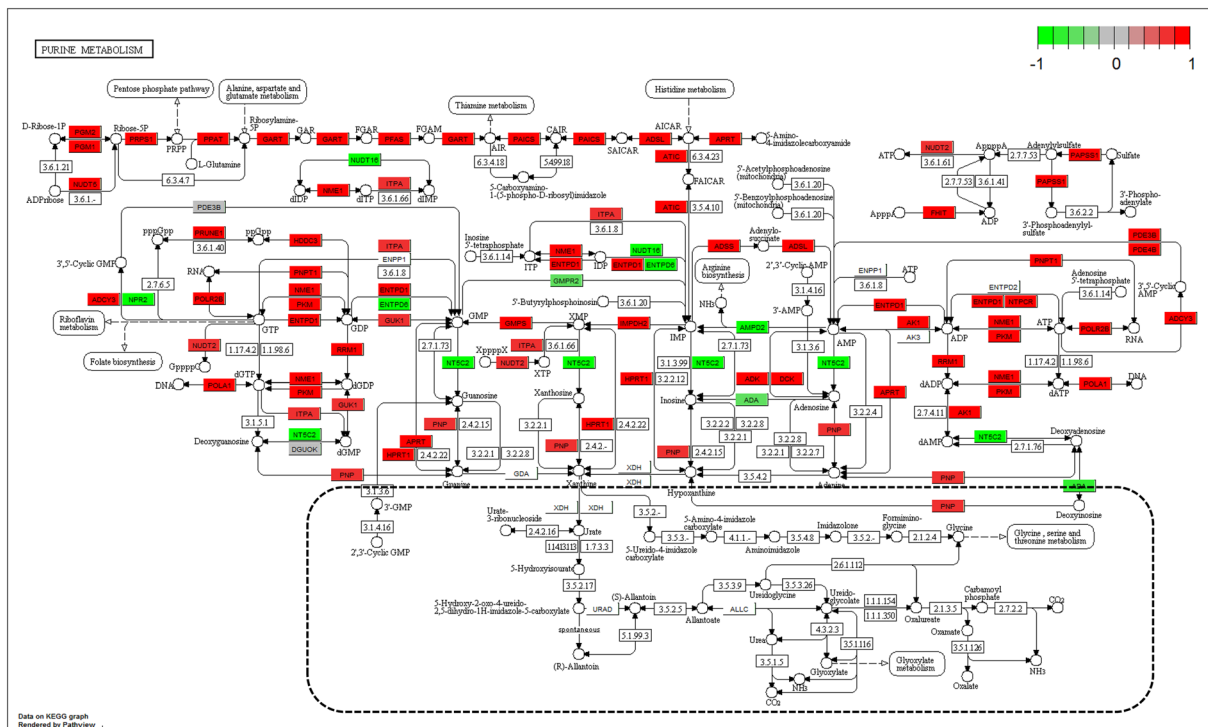
The p53 signaling pathway is a key factor that helps to conserve the stability of the genome. It regulates a wide variety of cellular processes including apoptosis, senescence, cell cycle arrest, differentiation and so on. The p53 signaling pathway is significantly changed after exposing to bryostatin (Fig. 8). Genes *CDK1*, *CDK2*, *CDK4*, *CCNE1*, *CCNB1* were up-regulated, which are contributes to the cell cycle arrest. Gene *CASP3* is up-regulated, which will lead to apoptosis. Therefore, the cell cycle arrest and apoptosis may have a potential role in the reactivation of HIV latency.

### Discussion

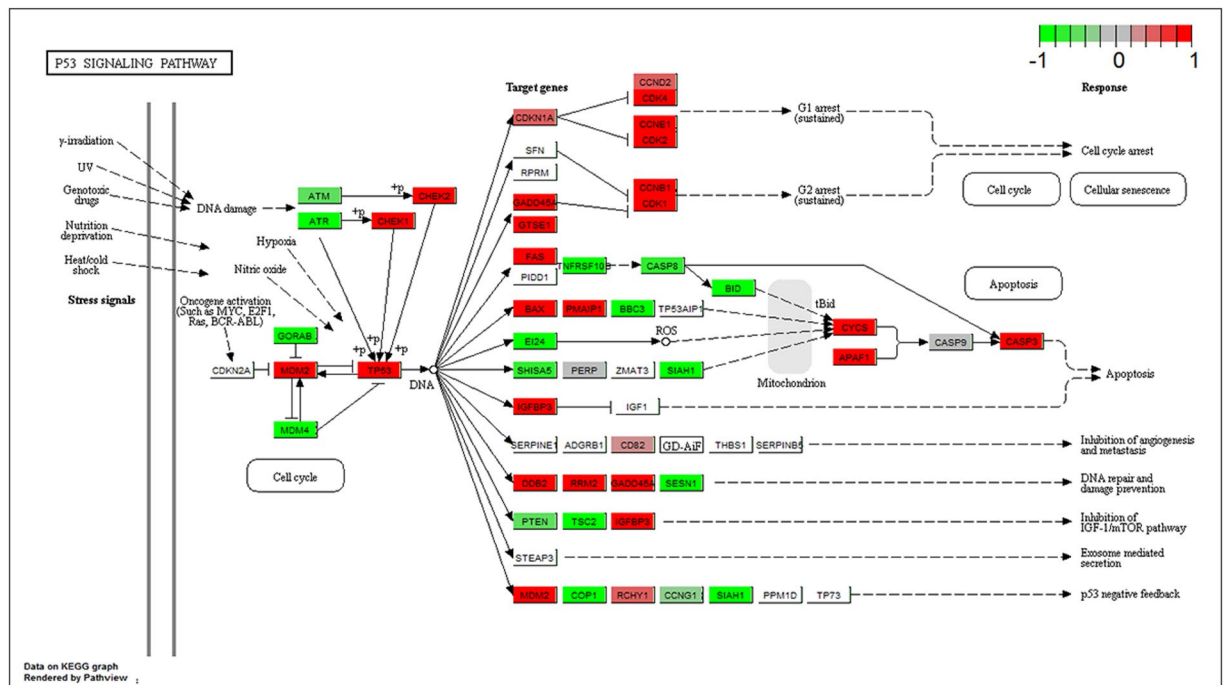
The latent reservoir is mostly found in resting memory CD4<sup>+</sup> T cells<sup>22,23</sup>. The most worrisome reservoir consists of latently infected resting memory CD4<sup>+</sup> T cells carrying integrated HIV-1 DNA. Bryostatin has shown significant potency to revert HIV latency in *ex vivo* experiment using patient cells, compared with other LRAs<sup>20</sup>. In this study, we focused to characterize the transcriptome in both the bryostatin treated and unstimulated HIV latent CD4<sup>+</sup> T cells. Bryostatin stimulation will lead to reactivation of HIV-latent CD4<sup>+</sup> T cells. To further confirm the change of HIV gene expression before and after reactivation, we checked the study that generated and provided the data<sup>24</sup>. The result showed that bryostatin significantly reactivated the latency cell compared with unstimulated condition (Fig. S1). And bryostatin has the highest efficiency at latency reversal among LRAs. Furthermore, as the importance of interleukins in the reactivation, we checked the gene expression for all interleukins and their receptors. We showed the genes that expressed in at least one sample (Fig. S2). The result indicates that the expression of interleukins and their receptors were heterogeneous. Only three genes IL7R, IL2RA and IL2RA were differential expressed between unstimulated and bryostatin treated samples. IL7R was down-regulated in the bryostatin treated samples. However, IL2RA and IL2RA were up-regulated in the bryostatin treated samples.



**Figure 6.** Gene expression profiles of pyrimidine metabolism pathway (hsa00240) in the bryostatin treated CD4<sup>+</sup> T cells. The red and green colors represent the log<sub>2</sub>(FC) of the corresponding gene in pyrimidine metabolism pathway. The dotted line circles the inactive part of this pathway<sup>50–52</sup>.



**Figure 7.** Gene expression profiles of purine metabolism pathway (hsa00230) in the bryostatin treated CD4<sup>+</sup> T cells. The red and green colors represent the log<sub>2</sub>(FC) of the corresponding gene in purine metabolism pathway. The dotted line circles the inactive part of this pathway<sup>50–52</sup>.



**Figure 8.** Gene expression profiles of p53 signaling pathway (hsa04115) in the bryostatin treated CD4<sup>+</sup> T cells. The red and green colors represent the log<sub>2</sub>(FC) of the corresponding gene in p53 signaling pathway<sup>50–52</sup>.

Most interleukins and their receptors did not show differential expression, which suggests interleukins and their receptors (including IL7R) may not affect ART efficiency in this study.

Accompanying the reactivation of HIV latency, we found extensive changes of gene expression. We identified a total of 597 DEGs, and most of DEGs are up-regulated. Among the top 30 DEGs, the genes *CXCR6*, *DNMT1* and *JOSD1* had studies on the HIV infection<sup>25–27</sup>, but had no reports on the HIV latency-reversal. Moreover, genes *CENPP*, *CENPU* and *CENPM*, as the component of the centromere complex (nucleosome-associated), were all from centromere protein (CENP) family and upregulated in the HIV latent CD4<sup>+</sup> T cells treated by bryostatin. Specifically, the gene *CENPP* was the component of CENPA (Histone H3-Like Centromeric Protein A)-CAD (nucleosome distal) complex, and the genes *CENPU/M* were the component of CENPA-NAC (nucleosome-associated) complex<sup>28</sup>. Cellular chromatin provides essential structure to the viral genome and is necessary for successful completion of the HIV life cycle. HIV can disrupt host cellular chromatin during heterochromatin integration and latent reactivation in unique centromere protein (CENPP, CENPU and CENPM) ways. These upregulated genes in the CD4<sup>+</sup> T cells stimulated by bryostatin may play an important role in the HIV latency-reversal.

The biochemical mechanisms involved in the induction of pro-viral transcription in latent HIV-infected CD4<sup>+</sup> T cells have not been fully defined. With the development of ARV therapy, it is important to understand how HIV infection changes the host cells pathways and metabolism. In our study, we found that pyrimidine metabolism, purine metabolism, and p53 signaling pathways are very closely related to HIV latency. Purines are key components of cellular energy systems (eg. ATP and NAD), signaling (eg. GTP, cAMP, and cGMP), and, along with pyrimidines, RNA and DNA production<sup>29,30</sup>. Therefore, pyrimidines and purines are crucial for nucleic acid biosynthesis, but also for lipid and carbohydrate. Previous studies revealed that pyrimidine and purine analogues were used as antiviral and immunosuppressive agents<sup>31,32</sup>. For example, arabinosylcytosine (ara-C), 5-fluorouracil (5-FU), and azidothymidine (AZT) have been used to target HIV reverse transcriptase<sup>32</sup>. In general, HIV infection will seriously impair the capability of T cells to synthesize the purine and pyrimidine ribonucleotide intermediates<sup>33</sup>. The shortage of these intermediates obstructs the HIV-infected cells to complete their cycle, thus turns the cell into latent state<sup>33</sup>. Therefore, bryostatin stimulates the latent cell and then results in up-regulation of pyrimidine and purine metabolism, which facilitates reactivation of the latent HIV infected cell.

p53 signaling pathway, a sensor for a broad range of cellular stresses, involved in several different aspects of cell cycle arrest, apoptosis, control of genome integrity, and DNA repair<sup>34</sup>. Cell cycle arrest and apoptosis are the most prominent outcomes of p53 activation<sup>35</sup>. In this study, p53 and its downstream genes (*CCNB1*, *CDK1*, *CCNE1*, *CKK2*, *CDK4*, and *CASP3*) were highly up-regulated (Fig. 8). The up-regulated genes *CCNB1* and *CDK1* will result in the sustained G2 arrest. Similarly, the up-regulated genes *CCNE1*, *CKK2*, and *CDK4* will lead to the sustained G1 arrest. The up-regulated gene *CASP3* will induce apoptosis. Therefore, we assume that bryostatin may lead to cell cycle arrest and apoptosis in the process of reactivation of latent HIV infected CD4<sup>+</sup> cell. Previous studies indicate that latent HIV-1 can sense the apoptosis of its host cell and responds by completing its replication cycle<sup>36,37</sup>. Another study revealed that G2 arrest of cell cycle induced HIV-1 transcriptional activation<sup>38</sup>. Together with previous studies, our results suggest that the reactivation of HIV latent cell, responding to bryostatin, may be partly due to the cell-cycle arrest and apoptosis induced by up-regulation of p53 signaling pathway.



It is worth noting that bryostatin, a potent activator of protein kinase C, may be involved in mitochondrial dysfunction<sup>39</sup>. Protein kinase C activation can induce mitochondrial dysfunction<sup>40–42</sup>. Protein p53 also can influence mitochondrial function<sup>43,44</sup>. Pyrimidine biosynthesis links mitochondrial respiration to the p53 pathway<sup>45</sup>. Therefore, mitochondrial dysfunction may be involved in the process that bryostatin reactivates the HIV-latent cell through up-regulation of pyrimidine and purine metabolism, and p53 signaling pathway.

In summary, we firstly revealed the extensive changes of gene expression in the process of HIV latent cells, responding to bryostatin. Hundreds of DEGs were identified and enriched both in the well-known and novel pathways. Secondly, we predicted 17 putative targets for bryostatin by two different methods. The 17 targets belong to multiple protein families including protein kinase, phospholipase A2, GPCR, peptidase M24A and so on. Finally, we discovered three novel pathways of pyrimidine metabolism, purine metabolism and p53 signaling important for the reactivation of HIV latency through functional network analysis of the bryostatin targets and DEGs. We assume that bryostatin reactivates the HIV-latent cell through up-regulation of pyrimidine and purine metabolism or through starting the cell-cycle arrest and apoptosis induced by up-regulation of p53 signaling pathway. Our study provides some novel insights into the role of bryostatin and its affected pathway in controlling HIV latency and reactivation.

## Methods

**Expression data collection and preprocessing.** The CD4<sup>+</sup> T cells expression data (GSE94150) was downloaded from Gene Expression Omnibus (GEO). The dataset was generated from the Illumina HiSeq. 2500 (Homo sapiens) and included 36 samples. The information for all samples was shown in Table S2. We described the key information of the samples here. All details can be found from <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE94150> and reference<sup>24</sup>. The CD4<sup>+</sup> T cells were isolated from a cohort of Florida HIV-infected subjects that have been on successful ART for >36 months with a viral load ≤50 copies/mL. The CD4<sup>+</sup> T cells were divided into three subsets namely central memory (T<sub>CM</sub>), transitional memory (T<sub>TM</sub>) and effector memory (T<sub>EM</sub>) CD4<sup>+</sup> T cells. The memory CD4<sup>+</sup> T cell T<sub>CM</sub>, T<sub>TM</sub>, and T<sub>EM</sub> subsets were stimulated for 24 h using different classes of agents including PKC activators (50 nM Bryostatin), gamma-c cytokines (10 ng/ml IL-15) and PMA+Ionomycin (100 ng/ml PMA and 1 μg/ml ionomycin). RNA-Seq was performed on the unstimulated and 24h-stimulated subsets. Unstimulated cells from the same donors were included as pre-activation baseline controls. Of 36 samples, 27 samples were stimulated with Bryostatin-1, IL-15 and PMA+Ionomycin, respectively. The other 9 samples were unstimulated. Our aim is to study transcriptional responses of CD4<sup>+</sup> T cells to the bryostatin-1 and 18 samples were retained (9 bryostatin-1 stimulated and 9 unstimulated CD4<sup>+</sup> T cells). The raw high-throughput mRNA sequencing reads was downloaded and analyzed by HISAT<sup>46</sup>, StringTie<sup>47</sup> and Ballgown<sup>48</sup> protocol. Firstly, the raw RNA-seq reads were mapped to the reference human genome using HISTA. Then, the alignments were passed to StingTie, which assembles and merges together all the gene structures found in any of the samples. In order to re-estimate the transcript abundance using the merged structures, the merged transcripts were fed back to StringTie again. Finally, Ballgown groups the all transcripts and abundances from StringTie by stimulated condition and determines which genes and transcripts are differentially expressed between conditions. The R v3.5.0 and Bioconductor library were used to analyze the expression data.

**KEGG pathways and biological processes analysis.** We used the functions `enrichKEGG` and `enrichGOR` in `clusterProfiler` R package<sup>49</sup> to perform KEGG pathways<sup>50–52</sup> and biological processes enrichment analysis. A pathway or process with a P value ≤ 0.05 was considered to be significantly enriched. The enriched pathways and processes were visualized using custom written R code. We selected top 10 significantly pathways affected by the compound bryostatin for further study. Therefore, we used the `pathview` R package<sup>53</sup> to display the expression profiles of the corresponding genes in these pathways. This package could provide the links between genes and pathways based on the KEGG pathways.

**Prediction of the targets of latency-reversal agent bryostatin.** In order to understand the action mechanism of LRA bryostatin, two methods of CGBVS<sup>54,55</sup> and 3NN<sup>56</sup> were used to predict the targets of bryostatin. The CGBVS is a chemical genomics-based virtual screening method for target prediction<sup>54</sup>. CGBVS predicts compound-protein interactions (CPIs) by using machine learn only based protein sequence rather than three-dimensional structures. The 3NN is a ligand-based similarity rankings method for target prediction<sup>56</sup>. 3NN method uses the 3 most similar scores (or their average value) between the query and the ligands sets in reference library to rank possible targets. We used venn diagram to show the number of target genes predicted by the two methods. Overlap targets were defined as the consistent targets of bryostatin.

**Functional interaction network analysis.** GeneMANIA<sup>57</sup> plug-in of Cytoscape<sup>58</sup> was applied to identify the functional interaction between the consistent targets of bryostatin and most DEGs. We used the consistent targets of bryostatin and most DEGs as input parameters. GeneMANIA (<http://www.genemania.org>)<sup>59</sup> can find other genes that are related to the set of input genes and produce a functional association network based on their relationships, such as pathways, co-expression, co-localization, genetic interaction, physical interaction, shared protein domains and so on, based on the published literature. Next, we used Cytoscape v3.2.1 to visualize the network.

## Data availability

All data analyzed during this study were described the methods and downloaded from <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE94150>.

Received: 26 June 2019; Accepted: 12 February 2020;

Published online: 26 February 2020

## References

- Churchill, M. J., Deeks, S. G., Margolis, D. M., Siliciano, R. F. & Swanstrom, R. HIV reservoirs: what, where and how to target them. *Nat. reviews. Microbiology* **14**, 55–60, <https://doi.org/10.1038/nrmicro.2015.5> (2016).
- Dahabieh, M. S., Battivelli, E. & Verdin, E. Understanding HIV latency: the road to an HIV cure. *Annu. Rev. Med.* **66**, 407–421, <https://doi.org/10.1146/annurev-med-092112-152941> (2015).
- Razooky, B. S., Pai, A., Aull, K., Rouzine, I. M. & Weinberger, L. S. A hardwired HIV latency program. *Cell* **160**, 990–1001, <https://doi.org/10.1016/j.cell.2015.02.009> (2015).
- Mohammadi, P. *et al.* Dynamics of HIV latency and reactivation in a primary CD4+ T cell model. *PLoS Pathog.* **10**, e1004156, <https://doi.org/10.1371/journal.ppat.1004156> (2014).
- Besnard, E. *et al.* Mechanism-The mTOR Complex Controls HIV Latency. *Cell host microbe* **20**, 785–797, <https://doi.org/10.1016/j.chom.2016.11.001> (2016).
- Nikolai, B. C. & Feng, Q. HIV Latency Gets a New Histone Mark. *Cell host microbe* **21**, 549–550, <https://doi.org/10.1016/j.chom.2017.04.012> (2017).
- Siliciano, R. F. & Greene, W. C. HIV latency. *Cold Spring Harb. Perspect. Med.* **1**, a007096, <https://doi.org/10.1101/cshperspect.a007096> (2011).
- Goldberg, A. D., Allis, C. D. & Bernstein, E. Epigenetics: a landscape takes shape. *Cell* **128**, 635–638, <https://doi.org/10.1016/j.cell.2007.02.006> (2007).
- Chen, L.-f., Fischle, W., Verdin, E. & Greene, W. C. Duration of Nuclear NF- $\kappa$ B Action Regulated by Reversible Acetylation.pdf. *Sci.* **293**, 1653–1657 (2001).
- Blazkova, J., Trejbalova, K., Gondois-Rey, F. & Halfon, P. CpG Methylation Controls Reactivation of HIV from Latency. *PLoS Pathog.* **5**, e1000554, <https://doi.org/10.1371/journal.ppat.1000554.g001> (2009).
- Shan, L. *et al.* Influence of host gene transcription level and orientation on HIV-1 latency in a primary-cell model. *J. virology* **85**, 5384–5393, <https://doi.org/10.1128/JVI.02536-10> (2011).
- Han, Y. *et al.* Orientation-dependent regulation of integrated HIV-1 expression by host gene transcriptional readthrough. *Cell host microbe* **4**, 134–146, <https://doi.org/10.1016/j.chom.2008.06.008> (2008).
- Darcis, G., Van Driessche, B. & Van Lint, C. HIV Latency: Should We Shock or Lock? *Trends immunology* **38**, 217–228, <https://doi.org/10.1016/j.it.2016.12.003> (2017).
- Williams, S. A. *et al.* Prostratin antagonizes HIV latency by activating NF-kappaB. *J. Biol. Chem.* **279**, 42008–42017, <https://doi.org/10.1074/jbc.M402124200> (2004).
- Archin, N. M. *et al.* Expression of latent HIV induced by the potent HDAC inhibitor suberoylanilide hydroxamic acid. *AIDS Res. Hum. Retroviruses* **25**, 207–212, <https://doi.org/10.1089/aid.2008.0191> (2009).
- Bouchat, S. *et al.* Histone methyltransferase inhibitors induce HIV-1 recovery in resting CD4(+) T cells from HIV-1-infected HAART-treated patients. *AIDS* **26**, 1473–1482, <https://doi.org/10.1097/QAD.0b013e32835535f5> (2012).
- Boehm, D. *et al.* BET bromodomain-targeting compounds reactivate HIV from latency via a Tat-independent mechanism. *Cell Cycle* **12**, 452–462, <https://doi.org/10.4161/cc.23309> (2013).
- Spina, C. A. *et al.* An in-depth comparison of latent HIV-1 reactivation in multiple cell model systems and resting CD4+ T cells from aviremic patients. *PLoS Pathog.* **9**, e1003834, <https://doi.org/10.1371/journal.ppat.1003834> (2013).
- Spivak, A. M. & Planelles, V. Novel Latency Reversal Agents for HIV-1 Cure. *Annu. Rev. Med.* **69**, 421–436, <https://doi.org/10.1146/annurev-med-052716-031710> (2018).
- Bullen, C. K., Laird, G. M., Durand, C. M., Siliciano, J. D. & Siliciano, R. F. New *ex vivo* approaches distinguish effective and ineffective single agents for reversing HIV-1 latency *in vivo*. *Nat. Med.* **20**, 425–429, <https://doi.org/10.1038/nm.3489> (2014).
- Schenone, M., Dancik, V., Wagner, B. K. & Clemons, P. A. Target identification and mechanism of action in chemical biology and drug discovery. *Nat. Chem. Biol.* **9**, 232–240, <https://doi.org/10.1038/nchembio.1199> (2013).
- Chomont, N. *et al.* HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nat. Med.* **15**, 893–900, <https://doi.org/10.1038/nm.1972> (2009).
- Eisele, E. & Siliciano, R. F. Redefining the viral reservoirs that prevent HIV-1 eradication. *Immun.* **37**, 377–388, <https://doi.org/10.1016/j.immuni.2012.08.010> (2012).
- Kulpa, D. A. *et al.* Differentiation to an effector memory phenotype potentiates HIV-1 latency reversal in CD4+ T cells. *Journal of virology*, <https://doi.org/10.1128/JVI.00969-19> (2019).
- Limou, S. *et al.* Multiple-cohort genetic association study reveals CXCR6 as a new chemokine receptor involved in long-term nonprogression to AIDS. *J. Infect. Dis.* **202**, 908–915, <https://doi.org/10.1086/655782> (2010).
- Rosca, A. *et al.* Immunoassay and molecular methods to investigate DNA methylation changes in peripheral blood mononuclear cells in HIV infected patients on cART. *J. Immunoass. immunochemistry* **38**, 299–307, <https://doi.org/10.1080/15321819.2016.1260587> (2017).
- Woodland, D. L. Theiler's Murine Encephalomyelitis Virus. *Viral immunology* **30**, 303, <https://doi.org/10.1089/vim.2017.29019.dlw> (2017).
- Izuta, H. *et al.* Comprehensive analysis of the ICEN (Interphase Centromere Complex) components enriched in the CENP-A chromatin of human cells. *Genes. cells: devoted Mol. Cell. mechanisms* **11**, 673–684, <https://doi.org/10.1111/j.1365-2443.2006.00969.x> (2006).
- Micheli, V. *et al.* Neurological disorders of purine and pyrimidine metabolism. *Curr. Top. Med. Chem.* **11**, 923–947 (2011).
- Nyhan, W. L. Disorders of purine and pyrimidine metabolism. *Mol. Genet. Metab.* **86**, 25–33, <https://doi.org/10.1016/j.ymgme.2005.07.027> (2005).
- Wiseman, A. C. Immunosuppressive Medications. *Clin. J. Am. Soc. Nephrology: CJASN* **11**, 332–343, <https://doi.org/10.2215/CJN.08570814> (2016).
- Okesli, A., Khosla, C. & Bassik, M. C. Human pyrimidine nucleotide biosynthesis as a target for antiviral chemotherapy. *Curr. Opin. Biotechnol.* **48**, 127–134, <https://doi.org/10.1016/j.copbio.2017.03.010> (2017).
- Bofill, M., Fairbanks, L. D., Ruckemann, K., Lipman, M. & Simmonds, H. A. T-lymphocytes from AIDS patients are unable to synthesize ribonucleotides de novo in response to mitogenic stimulation. Impaired pyrimidine responses are already evident at early stages of HIV-1 infection. *J. Biol. Chem.* **270**, 29690–29697 (1995).
- Levine, A. J. p53, the cellular gatekeeper for growth and division. *Cell* **88**, 323–331 (1997).
- Chen, J. The Cell-Cycle Arrest and Apoptotic Functions of p53 in Tumor Initiation and Progression. *Cold Spring Harb. Perspect. Med.* **6**, a026104, <https://doi.org/10.1101/cshperspect.a026104> (2016).
- Khan, S. Z., Hand, N. & Zeichner, S. L. Apoptosis-induced activation of HIV-1 in latently infected cell lines. *Retrovirology* **12**, 42, <https://doi.org/10.1186/s12977-015-0169-1> (2015).
- Badley, A. D., Sainski, A., Wightman, F. & Lewin, S. R. Altering cell death pathways as an approach to cure HIV infection. *Cell Death Dis.* **4**, e718, <https://doi.org/10.1038/cddis.2013.248> (2013).
- Thierry, S. *et al.* Cell cycle arrest in G2 induces human immunodeficiency virus type 1 transcriptional activation through histone acetylation and recruitment of CBP, NF-kappaB, and c-Jun to the long terminal repeat promoter. *J. virology* **78**, 12198–12206, <https://doi.org/10.1128/JVI.78.22.12198-12206.2004> (2004).

39. Wang, S., Wang, Z., Boise, L. H., Dent, P. & Grant, S. Bryostatins 1 enhances paclitaxel-induced mitochondrial dysfunction and apoptosis in human leukemia cells (U937) ectopically expressing Bcl-xL. *Leukemia* **13**, 1564–1573, <https://doi.org/10.1038/sj.leu.2401532> (1999).
40. Nowak, G., Takacsova-Bakajsova, D. & Megyesi, J. Deletion of protein kinase C-epsilon attenuates mitochondrial dysfunction and ameliorates ischemic renal injury. *Am. J. Physiol. Ren. Physiol* **312**, F109–F120, <https://doi.org/10.1152/ajprenal.00115.2016> (2017).
41. Nowak, G., Bakajsova, D. & Samarel, A. M. Protein kinase C-epsilon activation induces mitochondrial dysfunction and fragmentation in renal proximal tubules. *Am. J. Physiol. Ren. Physiol* **301**, F197–208, <https://doi.org/10.1152/ajprenal.00364.2010> (2011).
42. Nowak, G. Protein kinase C-alpha and ERK1/2 mediate mitochondrial dysfunction, decreases in active Na<sup>+</sup> transport, and cisplatin-induced apoptosis in renal cells. *J. Biol. Chem.* **277**, 43377–43388, <https://doi.org/10.1074/jbc.M206373200> (2002).
43. Dai, C. Q. *et al.* p53 and mitochondrial dysfunction: novel insight of neurodegenerative diseases. *J. Bioenerg. Biomembr.* **48**, 337–347, <https://doi.org/10.1007/s10863-016-9669-5> (2016).
44. Wang, D. B., Kinoshita, C., Kinoshita, Y. & Morrison, R. S. p53 and mitochondrial function in neurons. *Biochim. Biophys. Acta* **1842**, 1186–1197, <https://doi.org/10.1016/j.bbadis.2013.12.015> (2014).
45. Khutorenko, A. A. *et al.* Pyrimidine biosynthesis links mitochondrial respiration to the p53 pathway. *Proc. Natl Acad. Sci. USA* **107**, 12828–12833, <https://doi.org/10.1073/pnas.0910885107> (2010).
46. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements.pdf. *Nat. methods* **12**, 357, <https://doi.org/10.1038/nmeth.3317> (2015).
47. Pertea, M. *et al.* StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295, <https://doi.org/10.1038/nbt.3122> (2015).
48. Frazee, A. C. *et al.* Ballgown bridges the gap between transcriptome assembly and expression analysis. *Nat. Biotechnol.* **33**, 243–246, <https://doi.org/10.1038/nbt.3172> (2015).
49. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omic: a J. Integr. Biol.* **16**, 284–287, <https://doi.org/10.1089/omi.2011.0118> (2012).
50. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids Res.* **28**, 27–30, <https://doi.org/10.1093/nar/28.1.27> (2000).
51. Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K. & Tanabe, M. New approach for understanding genome variations in KEGG. *Nucleic acids Res.* **47**, D590–D595, <https://doi.org/10.1093/nar/gky962> (2019).
52. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**, 1947–1951, <https://doi.org/10.1002/pro.3715> (2019).
53. Luo, W. & Brouwer, C. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinforma.* **29**, 1830–1831, <https://doi.org/10.1093/bioinformatics/btt285> (2013).
54. Hamanaka, M. *et al.* CGBVS-DNN: Prediction of Compound-protein Interactions Based on Deep Learning. *Molecular informatics* **36**, <https://doi.org/10.1002/minf.201600045> (2017).
55. Yabuuchi, H. *et al.* Analysis of multiple compound-protein interactions reveals novel bioactive molecules. *Mol. Syst. Biol.* **7**, 472, <https://doi.org/10.1038/msb.2011.5> (2011).
56. Liu, X. *et al.* In Silico target fishing: addressing a “Big Data” problem by ligand-based similarity rankings with data fusion. *J. cheminformatics* **6**, 33, <https://doi.org/10.1186/1758-2946-6-33> (2014).
57. Zuberi, K. *et al.* GeneMANIA prediction server 2013 update. *Nucleic acids Res.* **41**, W115–122, <https://doi.org/10.1093/nar/gkt533> (2013).
58. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504, <https://doi.org/10.1101/gr.1239303> (2003).
59. Warde-Farley, D. *et al.* The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic acids Res.* **38**, W214–220, <https://doi.org/10.1093/nar/gkq537> (2010).

## Acknowledgements

This work was supported by the National Natural Science Foundation of China (31670945, 31401142), Yunnan Province Major Project for Science and Technology of China (2018IA044), the CAMS Innovation Fund for Medical Sciences (2017-12M-2-006), the start-up fund of Kunming university of science and technology (KKZ3201927005), and Yunnan Fundamental Research Projects (2019FB050). We thank our colleagues, Dr. Gong-Hua Li for helpful comments on the manuscript.

## Author contributions

M.S. and S.X.D. conceived the study and designed experiments. S.X.D., J.J.Z. and W.X.L. participated to acquisition and analysis of data. M.S., S.X.D., B.X.L., H.Z., Y.B.L., Y.L., K.F. jointly drafted the manuscript; S.X.D., J.J.Z. and W.X.L. jointly prepared the revised versions of the manuscript. All authors read and approved the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-60614-1>.

**Correspondence** and requests for materials should be addressed to M.S. or S.-X.D.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020