frontiers | Frontiers in Genetics

TYPE Brief Research Report
PUBLISHED 27 September 2024
DOI 10.3389/fgene.2024.1442539

OPEN ACCESS

EDITED BY
Zhongyi Yan,
Henan University, China

REVIEWED BY
Fang Yuan,
Henan University, China
Jing Liu,
Johns Hopkins University, United States
Huimin Li,
Henan University, China

*CORRESPONDENCE
Borbala Mifsud,
bmifsud@hbku.edu.qa

RECEIVED 02 June 2024
ACCEPTED 09 September 2024
PUBLISHED 27 September 2024

CITATION
Ahmed N, Cavattoni I, Villiers W, Cugno C, Deola S and Mifsud B (2024) Multi-omic analysis of longitudinal acute myeloid leukemia patient samples reveals potential prognostic markers linked to disease progression.
Front. Genet. 15:1442539.
doi: 10.3389/fgene.2024.1442539

Multi-omic analysis of longitudinal acute myeloid leukemia patient samples reveals potential prognostic markers linked to disease progression

Nisar Ahmed[1], Irene Cavattoni[2], William Villiers[1,3], Chiara Cugno[4], Sara Deola[4] and Borbala Mifsud[1,5]*

[1]College of Health and Life Sciences, Genomics and Precision Medicine, Hamad Bin Khalifa University, Doha, Qatar, [2]Hematology and Bone Marrow Transplant Unit, Ospedale Centrale Bolzano, Bolzano, Italy, [3]Department of Medical and Molecular Genetics, King's College London, London, United Kingdom, [4]Advanced Cell Therapy Core, Research, Sidra Medicine, Doha, Qatar, [5]William Harvey Research Institute, Queen Mary University London, London, United Kingdom

Relapse remains a determinant of treatment failure and contributes significantly to mortality in acute myeloid leukemia (AML) patients. Despite efforts to understand AML progression and relapse mechanisms, findings on acquired gene mutations in relapse vary, suggesting inherent genetic heterogeneity and emphasizing the role of epigenetic modifications. We conducted a multi-omic analysis using Omni-C, ATAC-seq, and RNA-seq on longitudinal samples from two adult AML patients at diagnosis and relapse. Herein, we characterized genetic and epigenetic changes in AML progression to elucidate the underlying mechanisms of relapse. Differential interaction analysis showed significant 3D chromatin landscape reorganization between relapse and diagnosis samples. Comparing global open chromatin profiles revealed that relapse samples had significantly fewer accessible chromatin regions than diagnosis samples. In addition, we discovered that relapse-related upregulation was achieved either by forming new active enhancer contacts or by losing interactions with poised enhancers/potential silencers. Altogether, our study highlights the impact of genetic and epigenetic changes on AML progression, underlining the importance of multi-omic approaches in understanding disease relapse mechanisms and guiding potential therapeutic interventions.

KEYWORDS

AML relapse, multi-omic analyses, chromatin reorganization, epigenetic modifications, Omni-C, ATAC-seq, RNA-seq

Introduction

Acute myeloid leukemia (AML) manifests as a complex disease marked by a multitude of genetic mutations and dysregulated gene expression profiles stemming from genetic and epigenetic alterations. These factors shape the trajectory of AML progression and confer resistance to therapeutic modalities. In general, AML occurs at any age, but it is the most prevalent form of acute leukemia in adults with a median age at diagnosis of 68 years and an estimated 20,380 diagnoses and 11,310 related deaths were projected for 2023 (Kishtagari and Levine, 2021; Siegel et al., 2023). In the past decade, extensive research has focused on

AML heterogeneity at disease onset, leading to improved classification (Döhner et al., 2022) and novel treatment agents (DiNardo and Cortes, 2016; Stein et al., 2017; Stone et al., 2017) that considerably help to achieve complete remission in most patients, however, the 5-year overall survival (OS) rates are still only at around 28% (Howlader N, et al., 2021). It is mainly due to the high relapse rate, as 40%–60% of patients relapse within 3 years and fail to respond to conventional chemotherapy regimen (Bejanyan et al., 2015; Karlsson et al., 2017; Verma et al., 2010). Unfortunately, most individuals who relapse ultimately die from the disease (Schlenk et al., 2018). Prognosis in case of relapse is typically more unfavorable, especially when the recurrence occurs within a year after the initial remission (Rasche et al., 2021). Furthermore, relapse is a primary factor contributing to treatment failure (Döhner et al., 2015).

Recently, several investigations have leveraged next-generation sequencing (NGS) to identify gene mutations specific to relapse in certain AML subgroups, aiming to elucidate the disease's course. However, there is a considerable genetic heterogeneity among different relapsed patients (Jan et al., 2012; Metzeler et al., 2016; Papaemmanuil et al., 2016), even when focusing on specific AML subgroups (Ahn et al., 2018; Gröschel et al., 2015; Wang et al., 2016). This diversity hinders the discovery of consistent relapse-specific signatures. Additionally, there are only a few studies that compare the mutational profiles of diagnosis and relapse samples in AML patients, and they focus on coding mutations (Farrar et al., 2016; Masetti et al., 2016), despite growing evidence that non-coding mutations in regulatory elements or structural variants altering enhancer usage can also drive oncogenesis (Northcott et al., 2014; Zhu et al., 2020). This underscores the significance of epigenetic changes in the course of the disease and emphasizes the necessity of considering the intrinsic and extrinsic disease heterogeneity (Levin et al., 2021; Schwenger and Steidl, 2021; Vicente-Dueñas et al., 2018). Furthermore, there is a paucity of comprehensive molecular characterization of longitudinal AML samples, including diagnosis and relapse pairs. Therefore, we aim to assess the contribution of epigenetic factors, such as active regulatory elements and their long-range chromatin interactions, in conjunction with gene expression profiles across longitudinal samples, elucidating their role in the progression of AML.

To achieve this, we have profiled the long-range chromatin interactions (using Omni-C), open chromatin regions (using ATAC-seq), and gene expression (using RNA-seq) of two matched adult AML patients at diagnosis and relapse. We found significant alterations in the 3D chromatin landscape and accessible chromatin regions between relapse and diagnosis samples. Additionally, upregulated genes in relapse showed enrichment for H3K27me3 in distal regions of diagnosis-specific interactions, indicating loss of potential silencer connections during relapse.

Methods

Patient samples

We received live-frozen mononuclear cells from peripheral blood of adult AML diagnosis and relapse paired samples from Bolzano General Hospital, Italy. Each sample contained >50% blast cells Supplementary Table S1. The project was reviewed by the Hamad Bin Khalifa University Institutional Review Board and approved under protocol #QBRI-IRB 2020-02–017.

Omni-C library preparation

We used the Omni-C kit (Dovetail Genomics), which is a sequence-independent endonuclease-based proximity-ligation protocol. Briefly, 1×10^6 live bone marrow cells were fixed in DSG (disuccinimidyl glutarate), a non-cleavable and membrane-permeable protein-protein crosslinker, followed by formaldehyde to reversibly crosslink *in vivo* DNA-protein interactions. The fixed cells were treated with DNase I to digest chromatin. Next, for the proximity ligation, the chromatin ends were polished, and biotin-tagged bridges were used to create chimeric molecules. The crosslink of lysate was reversed, and the purified DNA was used for NGS library preparation. Finally, the library was enriched for ligation-containing chimeric molecules. The Omni-C libraries were sequenced on an average of 14x coverage on the Illumina HiSeq X Ten system with 151-base paired-end reads (>300M reads).

ATAC-seq library preparation

One of our objectives was to interrogate active chromatin regions through an assessment of chromatin accessibility, which was evaluated using ATAC-seq, a method known for its capability to delineate regions of open chromatin. The Active Motif ATAC-seq kit was used to perform ATAC-seq on living cells in accordance with the Omni-ATAC-seq protocol described by (Corces et al., 2017). For ATAC-seq, cryopreserved bone marrow samples were slowly thawed using IMDM supplemented with 10% FBS and DNase. Viability was calculated under a hemocytometer with trypan blue–samples with a viability <80% were subjected to dead cell sorting using MACS dead cell removal kit (cat: 130-090–101). 1×10^5 cells were taken forward for ATAC-seq.

Briefly, for sample preparation, 1×10^5 cells were first pelleted and washed with ice-cold PBS. Subsequently, the cells were re-suspended in an ice-cold ATAC-Lysis buffer. Next, for tagmentation the isolated nuclei were incubated at 37°C for precisely 30 min while being shaken at 800 rpm in a transposition mixture containing 100 nM final transposase. For DNA purification Zymo DNA Clean and Concentrator-5 Kit was used. Following that, PCR amplifications of tagmented libraries were performed using 10 cycles of PCR and DNA was extracted using 60 µL SPRI beads for size selection. To assess size distribution, PCR-amplified libraries were analyzed with Bioanalyzer. Finally, libraries were sequenced on the Illumina HiSeq 4,000 platform to ~50 million paired end 100bp reads.

RNA-seq library preparation

Total RNA was purified from 5×10^5 live bone marrow cells using a Qiagen RNeasy Plus Micro kit. Briefly, the cells were disrupted and homogenized using RLT buffer. Genomic DNA eliminator spin columns were used to remove the DNA, and

RNeasy MinElute spin columns were used to purify RNA. Total RNA-seq libraries were prepared using TruSeq Stranded Total RNA Library Prep Gold Kit. The prepared libraries were sequenced using the Illumina Nextseq platform to ~60–50 million paired-end 101bp reads.

Omni-C data pre-processing

Omni-C libraries were processed through an in-house pipeline. The quality control (QC) of sequencing of Omni-C libraries was performed on fastq files using FastQC (Andrews et al., 2010). The adaptor sequences from the reads were trimmed using TrimGalore. The trimmed read pairs were aligned using BWA-MEM (version 0.7.17 or higher) to the GRCh38 version of the human genome. To find ligation junctions in Omni-C libraries, the "pairtools parse" module was used by setting MAPQ greater or equal to 40 and walks-policy as 5unique. The pairtools pipeline records the strand of each paired read and the outermost (5′) aligned base pair into a "pairsam" file upon identification of a ligation event in the alignment file. The pairsam format records Hi-C pair information along with SAM entries, which was then sorted using "pairtools sort". The "pairtools dedup" command was used to remove PCR duplicates from sorted pairsam files and here we also produced the statistics of the library by using the flag "–output-stats". Finally, the deduplicated pairsam files were then used to create two different files, such as pairs file and BAM file, which can be used for further downstream processing.

Omni-C library QC and complexity

To check the quality of Omni-C libraries, we have used the stats files calculated by "pairtools dedup", which contains information on total reads, mapped reads, duplicate reads, and total read pairs. In addition, we used the get_qc.py pipeline from Dovetail Genomics to summarize these stats in percentage and absolute values. We have checked the complexity of Omni-C libraries using the lc_extrap utility of the preseq package from Smith lab (github.com/smithlabcode/preseq), which aims to predict the complexity of sequencing libraries.

Contact matrix

The contact maps, which are compressed and sparse formats, are produced from the pairs files using Juicer tools (Durand et al., 2016). The pairs files were first converted into HiC files, which are highly compressed binary representations of the contact matrices using "pre" command of Juicer tools. The HiC contact matrices were finally visualized using Juicebox (J. T. Robinson et al., 2018).

Differential interaction analysis

The systematic biases such as those arising from enzyme digestion, DNA ligation, and PCR amplification from Omni-C libraries were corrected using the HiCorr pipeline (Lu et al., 2020). To prepare the input, the BAM files were sorted according to co-ordinates and read pairs were then mapped to select the cis and trans read pairs. The HiCorr outputs were then used to apply a deep learning-based tool called DeepLoop (Zhang et al., 2022) to perform loop signal enhancement. We have used the pre-built models from DeepLoop to improve the sensitivity, robustness, and quantitation of Omni-C loops and output chromatin loop strength.

After that, we have created a count matrix ($M_{(i,\ j)}$) where each row *(i)* represents a chromatin loop and column *(j)* represents a sample and it was populated with the loop strengths from DeepLoop. The count matrix was normalized using the R Bioconductor package edgeR (M. D. Robinson et al., 2010). The normalized counts were used to calculate the standard deviation, and we selected the top 100,000 most variable interactions based on standard deviation. Finally, the matrix with only the most variable interactions was used to perform differential interaction analysis using the R Bioconductor package limma (Ritchie et al., 2015).

Omni-C downstream analysis

Exploratory analysis was performed in R version 4.3.1 using the Bioconductor package GenomicRanges (Lawrence et al., 2013). All interaction landscapes were visualized in the WashU Epigenome Browser (Li et al., 2022).

ATAC-seq library pre-processing

Briefly, all libraries underwent FastQC testing (Andrews et al., 2010) to evaluate the library quality and make sure that each library is free of significant problems like low read quality or adapter contamination. Next, we used Trimmomatic (Bolger et al., 2014) with default parameters to filter low-quality reads, and Truseq adaptors were trimmed off from the reads. Subsequently, these reads were aligned to the hg38 version of the human genome using Bowtie2 aligner, which created SAM files that were converted to BAM files using samtools. After that, mitochondrial, duplicate, and blacklisted reads were removed using samtools and bedtools. Reads were shifted to adjust for tn5 binding using the alignmentSieve tool. Finally, peaks were then called on the final processed BAM files using the callpeak command with BAMPE of the MACS2 peak calling algorithm.

ATAC-seq downstream analysis

MACS2 peaks were further used for downstream processing. Peaks were assigned to genomic elements using BioMart (Durinck et al., 2005) and TxDb.Hsapiens.UCSC.hg38. knownGene R Bioconductor packages (Carlson, et al., 2015). All exploratory analysis was performed in R version 4.3.1 using the GenomicRanges Bioconductor package (Lawrence et al., 2013).

RNA-sequencing pre-processing

The QC of RNA-seq libraries was performed on fastq files using the FastQC tool (Andrews et al., 2010). The adaptor sequences from

the reads were trimmed using TrimGalore. The trimmed reads were aligned to the hg38 genome using STAR and RSEM was used to calculate the expression values as expected counts from the aligned RNA-seq data. The count matrix was normalized using the R Bioconductor package edgeR (M. D. Robinson et al., 2010). Finally, the normalized counts were used to analyze differential gene expression using the R Bioconductor package limma (Ritchie et al., 2015). Significant upregulated and downregulated genes with adjusted *p*-value <0.1 were selected based on log2 fold change >1 and log2 fold change < −1, respectively.

RNA-seq downstream analysis

The normalized reads were used for all exploratory analysis and plots were generated using custom code in R version 4.3.1. Volcano plots were created using the R Bioconductor package ggplot2.

Multi-omic data integration

Significant differential Omni-C interactions were used to identify relapse-specific and diagnosis-specific anchors in each sample and subsequently, we specified diagnosis-specific ATAC-seq peaks by taking the overlap of two diagnosis samples and relapse-specific ATAC-seq peaks by taking the overlap of two relapse samples. Further, we removed the common peaks between diagnosis-specific and relapse-specific ATAC-seq peaks. Next, relapse-specific and diagnosis-specific Omni-C anchors and ATAC-seq peaks were overlapped using the Bioconductor package GenomicRanges (Lawrence et al., 2013) to investigate which differentially interacting regions are accessible. Finally, we assigned genes to these regions using BioMart (Durinck et al., 2005).

Furthermore, we investigated which differentially expressed genes are linked to differentially accessible DNA in differentially interacting regions (20 Kb). Then, we calculated enrichment for differentially expressed genes in relapse and diagnosis-specific interactions in comparison with 10 random, size-matched sets of genes from the whole genome. Finally, the distal or other end of the interactions of upregulated genes that have diagnosis-specific Omni-C interactions and ATAC-seq peaks were overlapped with publicly available H3K27me3 ChIP-seq data (accession number GSM4565992) from $CD34^{+}$ common myeloid progenitor cells, to test whether their upregulation can be attributed to the loss of a silencer contact. For these analyses, the R Bioconductor package GenomicRanges (Lawrence et al., 2013) was used.

Motif enrichment

For the differential enrichment of transcription factor binding motifs within differentially accessible regions found in differentially interacting regions the Homer (Hypergeometric Optimization of Motif EnRichment) motif suite of tools was used (Heinz et al., 2010). The parameters findMotifsGenome.pl with -size −200,200 to centre peaks to a 400bp region and -bg to set background of peaks. Motifs with FDR<0.01 were considered significantly enriched. Relapse-specific anchors and peaks were used as target sequences, and diagnosis-specific anchors and peaks were used as background sequences.

Results

Overview of multi-omics assays in diagnosis and relapse AML

We integrated changes of the 3D genome structure, chromatin accessibility, and gene expression to decipher the molecular changes that occur in AML at relapse compared to at the time of diagnosis in two initial diagnoses and relapse AML sample pairs (Supplementary Table S1). We explored long-range regulatory interaction patterns using Omni-C, active regulatory elements using ATAC-seq, and transcriptional patterns through RNA-seq (Figure 1A). In addition, we used Omni-C data to create genome-wide contact maps and assess the spectrum of large chromosomal changes in these AML samples. The contact maps identified that all of the samples had a t(9; 11) or *KMT2A::MLLT3* translocation. This confirmed the initial diagnosis of one patient; however, this translocation was undetected by cytogenetics in the other patient. The KMT2A::MLLT3 subtype of AML is shown to have a poor/intermediate prognosis, and the current mechanistic understanding of KMT2A-rearrangement (KMT2Ar) prognosis has not fully translated into therapeutic success due to the complexity of genomic events contributing to the disease (Krivtsov and Armstrong, 2007; Liedtke and Cleary, 2009; Meyer et al., 2018). We additionally observed within our Omni-C data that relapse samples had gained extra chromosomal abnormalities, for example, chromosome eight duplication, t(3; 5), and t(1; X) (Figure 1B; Supplementary Figure S1A–D). Such abnormalities are common occurrences at relapse in AML (Kern et al., 2002).

Differential interactions and differentially accessible regions in relapse *versus* diagnosis samples

Next, we asked whether there are any regions that are differentially interacting in the genome distinguishing relapse from diagnosis samples. While we found higher intra-patient similarity than intra-status (Figure 2A), we could detect common diagnosis and relapse-specific interactions. By comparing relapse *versus* diagnosis samples, we found 8,202 significantly differential interactions (p-adj <0.1), where 5,262 interactions were defined as diagnosis-specific, and 2,940 interactions were found to be relapse-specific (Figure 2B). This highlights distinctive interaction patterns between diagnosis and relapse timepoints that are consistent across patients. Next, we compared the chromatin accessibility profile of the diagnosis and relapse pairs using ATAC-seq. We estimated the similarity between relapse and diagnosis samples using PCA based on the global open chromatin profile, which again showed higher inter-patient variability than inter-state variability (Figure 2C). We performed an overlap analysis between two diagnosis samples and two relapse samples, where we took the common diagnosis (20,197) and common relapse (10,914) open chromatin regions. Subsequently, by overlapping common diagnoses and relapse peaks, we identified 16,363 consistent diagnosis-specific open

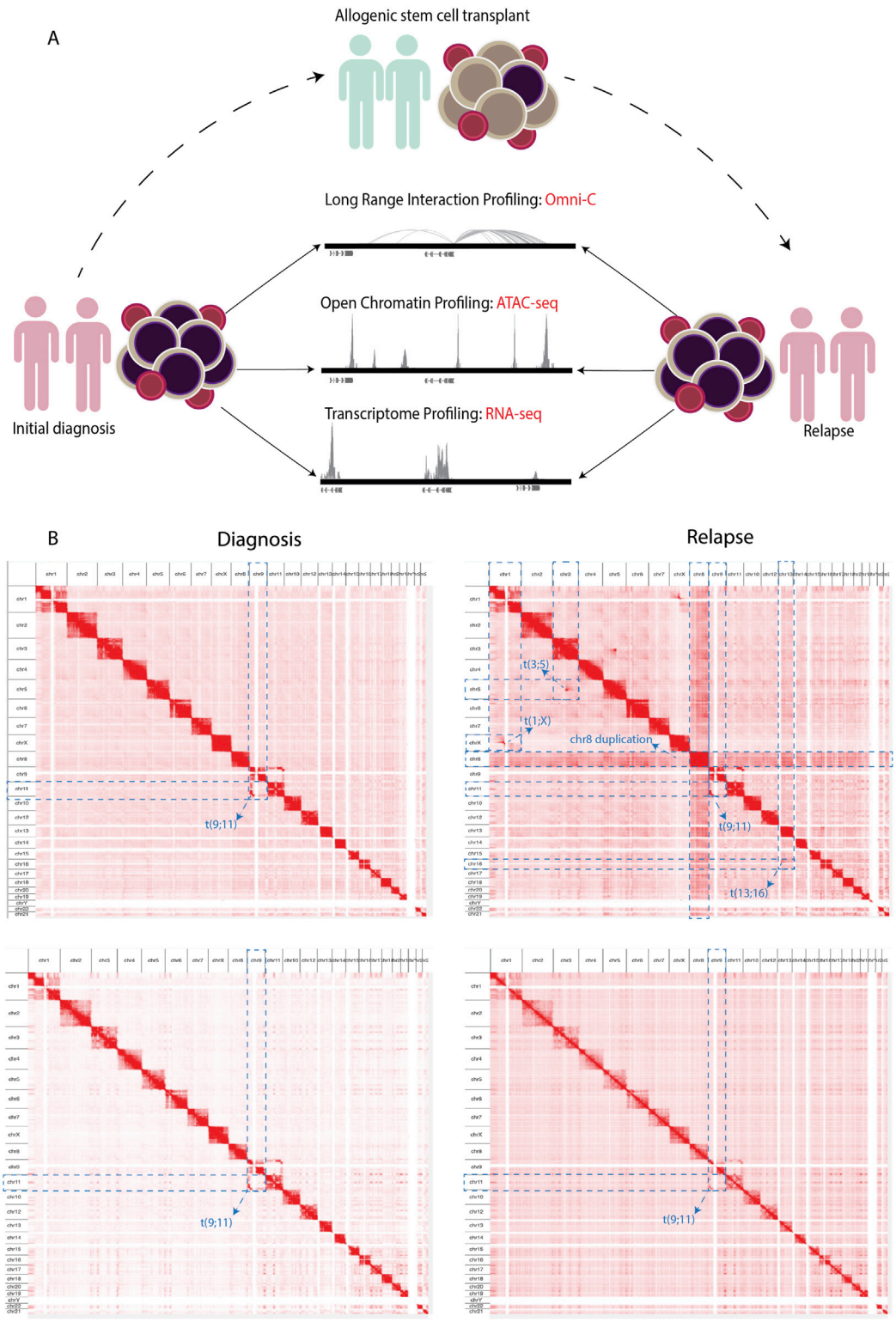

FIGURE 1
(A) Schematics showing the experimental workflow where PBMCs obtained from adult AML samples at initial diagnosis underwent NGS library preparation for Omni-C, ATAC-seq, and RNA-seq analyses. Subsequently, the same NGS libraries were prepared from samples upon relapse following allogenic stem cell transplantation. **(B)** Contact maps derived from the Omni-C dataset, depicting the chromosomal architecture of diagnosis samples (*left*) compared with relapse samples (*right*). The blue dotted lines show the chromosomal abnormalities each sample contains genome wide.

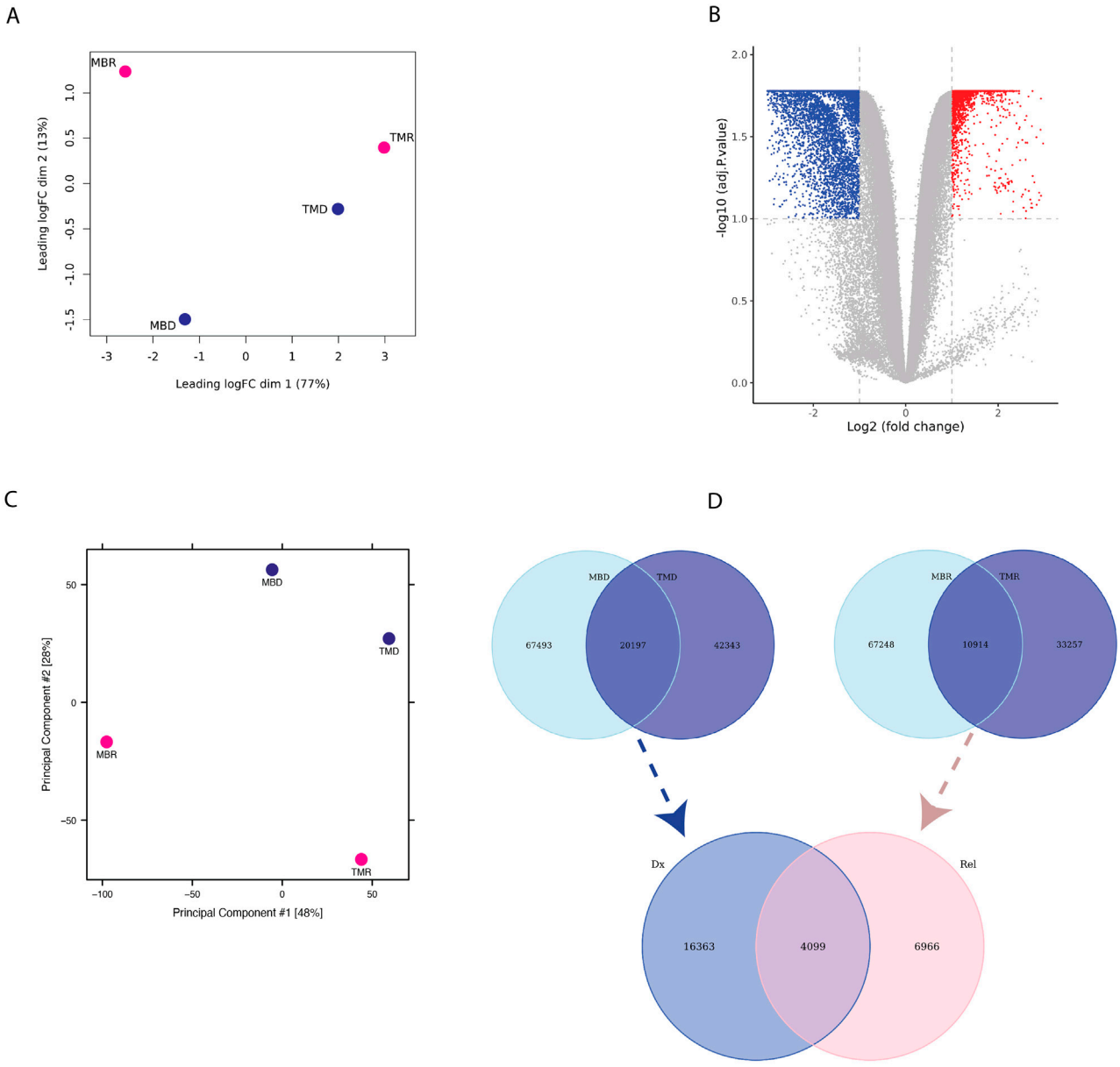

FIGURE 2
(A) PCA plot based on limma logfold change (logFC) of Omni-C data **(B)** Volcano-plot showing differential interaction of relapse *versus* diagnosis interactions based on Omni-C datasets where each dot represents an interaction. Blue represents significant diagnosis-specific interactions (5,262), and red represents significant relapse-specific interactions (2,940), whereas grey dots represent non-significantly differential interactions. **(C)** PCA plot based on ATAC-seq normalized peak intensity. **(D)** Venn diagram representing the overlap of diagnosis-specific ATAC-seq peaks (*upper left*), relapse-specific ATAC-seq peaks (*upper right*) between patients and of consistent diagnosis-specific and relapse-specific ATAC-seq peaks (*bottom*).

chromatin regions and 6,966 relapse-specific regions in addition to 4,099 regions that were active in all samples (Figure 2D).

Integration of differential chromatin interaction and accessibility signatures

We integrated the Omni-C anchors with ATAC-seq peaks to investigate which differential interacting regions are also differentially accessible (Supplementary Figure S2A). We found 150 unique relapse-specific anchors that were differentially accessible in relapse samples, which were annotated with 107 genes. GO-term analysis of these genes reveals the biological processes and pathways in which these genes were enriched, such as regulation of canonical Wnt signaling pathway, negative regulation of cell differentiation, Notch signaling pathway, and AML (Figure 3A). Next, we investigated if relapse-specific peaks in relapse-specific anchors are associated with different transcription factors (TFs) compared to diagnosis-specific ones by performing differential motif discovery on these two sets. We noted that a liver X receptor beta (LXRb) was enriched in relapse-specific regions (Figure 3B). Finally, we explored whether these differential

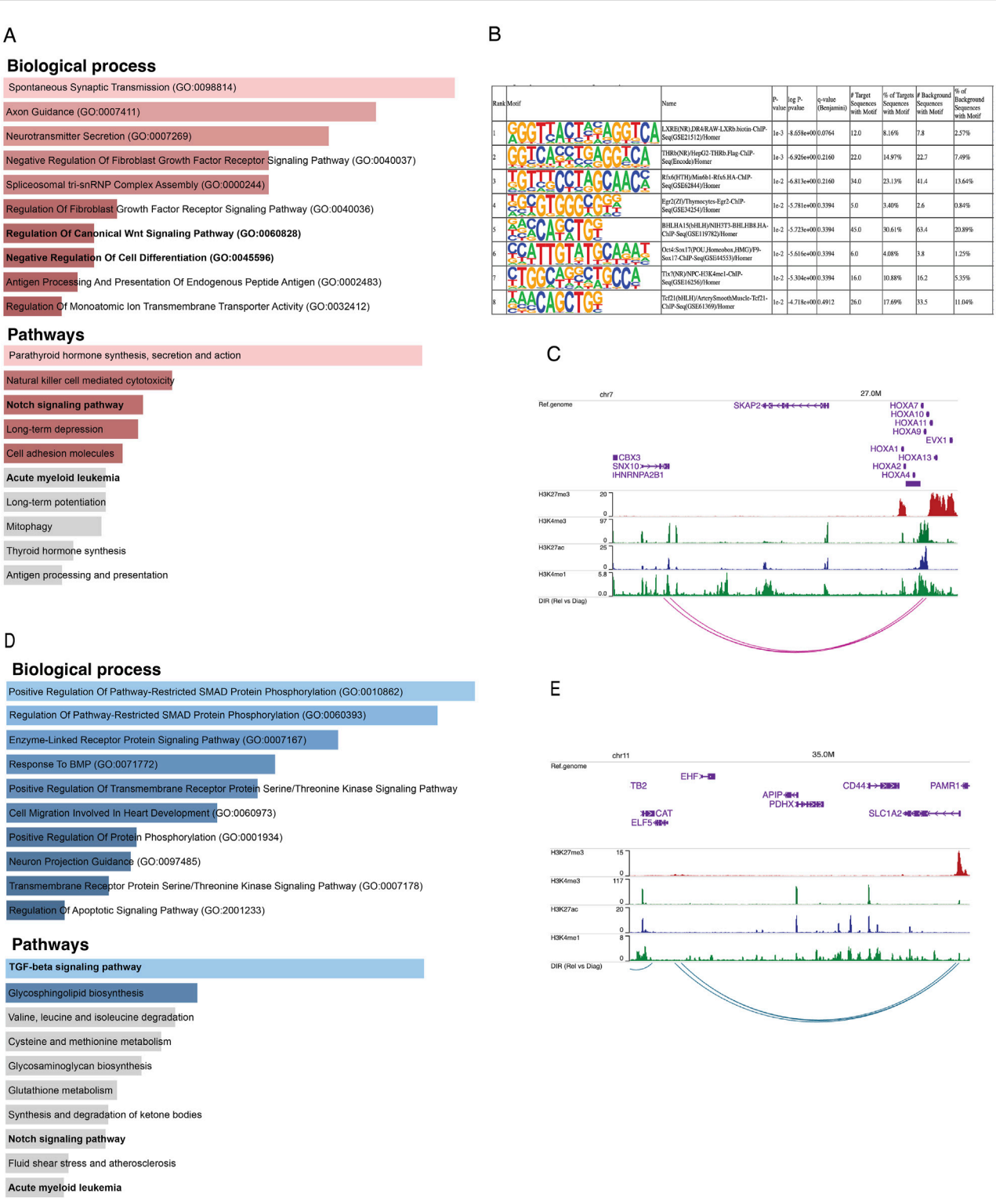

FIGURE 3
(A) Bar plots representing the enrichment of biological processes and pathways based on genes annotated to consistent relapse-specific interactions and peaks. **(B)** The top eight motifs enriched in relapse-specific Omni-C anchors and ATAC-seq peaks compared to diagnosis-specific anchors and peaks. **(C)** Example of a relapse-specific Omni-C interaction and ATAC-seq peaks where the *HOXA9* promoter region interacts with a distal active enhancer at the downstream region of the *SNX10* gene. **(D)** Bar plots representing the enrichment of biological processes and pathways based on genes annotated to diagnosis-specific interactions and peaks. **(E)** Example of a diagnosis-specific interaction and ATAC-seq peaks at the promoter region of *ELF5* interacting with a poised enhancer within *SLC1A2*. Arcs represent long-range chromatin interactions, active enhancers are marked by H3K27ac and H3K4me1, and poised enhancers are marked by H3K27me3 and H3K4me1.

10.3389/fgene.2024.1442539

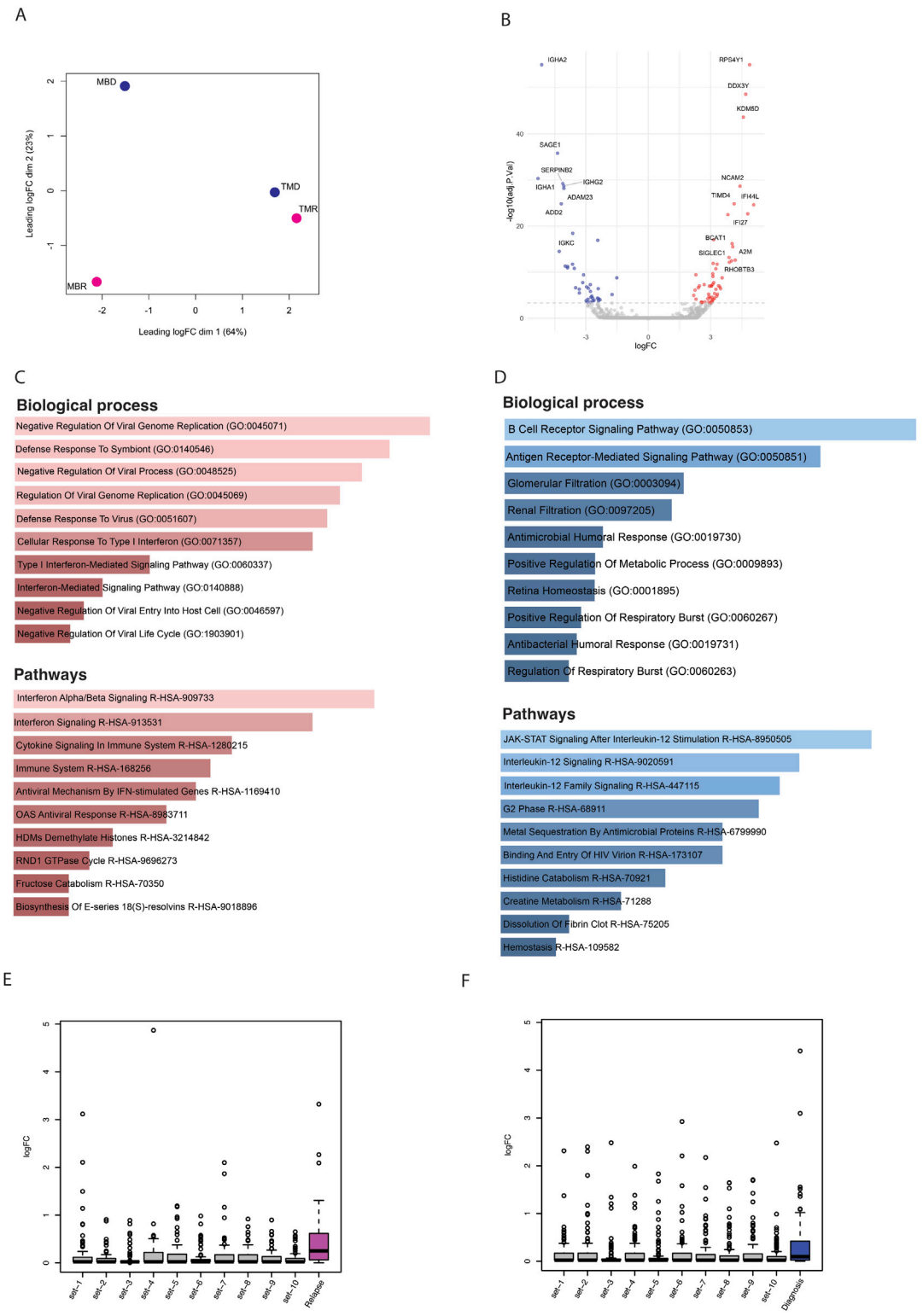

FIGURE 4
(A) PCA plot based on logfold change (logFC) of gene expression level. **(B)** Volcano plot showing significantly (p-adj <0.1) differentially expressed genes; here blue dots are significantly downregulated genes, red dots are significantly upregulated, and grey are nonsignificant genes. **(C)** Bar plots showing the enrichment of biological processes and pathways based on upregulated genes. **(D)** Bar plots showing enrichment of biological processes and pathways based on downregulated genes. **(E)** Box plots representing the logfold change of upregulated genes that overlapped with relapse-specific interactions and peaks (pink) and of 10 different random sets of upregulated genes. **(F)** Box plots representing the logfold change of upregulated genes that overlapped with diagnosis-specific interactions and peaks (blue) and of 10 different random sets of upregulated genes.

interactions are anchored at active or poised enhancers. For instance, *HOXA9*, which is involved in cell differentiation and upon dysregulation, contributes to leukemogenesis, was found to form a relapse-specific long-range interaction with an active enhancer at the 3′UTR of the *SNX10* gene, marked by H3K4me1/3 and H3K27ac (Figure 3C). In addition, we found KMT2Ar subtype-specific patterns at known KMT2A target genes, such as *UBE2J1* and *PARP8*, in the loop anchors (Supplementary Figures S2B, S2C). Overexpression of these genes contributes to the blockage of normal hematopoietic differentiation and promotes leukemogenesis.

On the other hand, we found 312 unique diagnosis-specific anchors that were differentially accessible and in proximity to these peaks, 154 genes were present. Further, these genes were enriched in biological pathways such as TGF-beta and Notch signaling (Figure 3D). Examples of genes with diagnosis-specific interactions include *ELF5*, which is associated with epithelial cell function and has implications in cancer. Its promoter formed a diagnosis-specific contact with a poised enhancer marked by H3K27me3 and H3K4me1, likely contributing to its low expression in diagnosis samples (Figure 3E). Additionally, a TGF-beta pathway member, *BMP8A*, also interacted with a poised enhancer. This gene plays critical roles in various cellular processes, such as tissue differentiation, cell proliferation, and apoptosis. *ST8SIA1*, which is involved in cell signaling, recognition, and adhesion, was interacting with an active enhancer near the 3′UTR of *CMAS*. Another gene with a diagnosis-specific active enhancer contact is *HPSE2*, which has a role in tumor microenvironment dynamics and cancer progression. Its promoter is interacting with an active enhancer present in the proximity of its own 3′UTR Supplementary Figure S2D–F).

Analysis of differential gene expression dynamics

Finally, we explored the gene expression profile of these longitudinal samples using RNA-seq. Similarly to what we found using the other omics methods, samples do not cluster together based on disease status, i.e., relapse and diagnosis (Figure 4A). Nonetheless, we asked whether there are differentially expressed genes between relapse *versus* diagnosis samples. Using limma-voom, we found overall, 95 significantly dysregulated genes that were selected using p. adj <0.1 and absolute log2FC > 1 (Figure 4B). These genes were enriched in pathways like B-cell receptor signaling pathway and immune and defense mechanism-related biological processes, as well as AML (Supplementary Figure S2G). Further, out of 95 significant dysregulated genes identified by comparing relapse with diagnosis, there were 55 upregulated genes (Supplementary Table S2) and 40 downregulated genes (Supplementary Table S3). These upregulated genes were enriched in defense response and interferon and cytokines signaling (Figure 4C), whereas downregulated genes were enriched in B-cell receptor, Jak-STAT, and interleukin signaling pathways (Figure 4D).

Multi-omic data integration exploration

Finally, we integrated all three omic datasets together to identify consistently differentially regulated genes across relapse and diagnostic timepoints. While there were only a handful of genes that showed significant differences, we noted that in general, genes that were associated with relapse-specific interactions and open chromatin regions showed higher levels of upregulation than those genes associated with diagnosis-specific interactions and open chromatin regions, as well as random genes from the genome, indicating that relapse-specific interactions are often new active enhancer contacts (Figures 4E, F). Some of the genes that were associated with diagnosis specific interactions and open chromatin regions were also upregulated in relapse samples. We hypothesized that these upregulated genes may have initially interacted with silencers/poised enhancers, but upon relapse, this interaction is lost (Supplementary Figure S3A). We found 13 genes (*ST8SIA1*, *IVD*, *ACAN*, *KIF5C*, *ADAMTS5*, *CFAP299*, *PITX2*, *KLHL31*, *RBPMS*, *SNAI2*, *SOX17*, *CSMD3*, and *DMRTA1*) where the diagnosis-specific interaction linked them to regions marked by H3K27me3, which is a significantly higher proportion than expected by chance (p-value = 9.99×10^{-5}) (Supplementary Figures S3B, C). In contrast, genes associated with either relapse-specific or diagnosis-specific interactions and open chromatin regions showed a similar level of downregulation compared to random genes from the genome (Supplementary Figures S3D, E). In an independent study, we found that out of these 13 genes, nine genes lost chromatin accessibility upon relapse (Nuno et al., 2024) (Supplementary Figure S4).

There were five genes that showed significant differences in all three analyses. Our findings revealed that among genes that were upregulated in relapse, *SDC2* and *CD70* were in consistent relapse-specific anchors and peaks, while *NCAM2* and *IFI44* were in consistent diagnosis-specific anchors and peaks. These genes have been shown to have some link with different types of cancers. For example, the SDC2 (syndecan-2) protein functions as an integral membrane protein and participates in cell proliferation, cell migration and cell-matrix interactions via its receptor for extracellular matrix proteins, and altered *SDC2* expression has been detected in several different tumor types (Akl et al., 2015; Canarte et al., 2023). Among the downregulated genes in relapse, *TSPYL5*, which is a *TP53* suppressor via its interaction with USP7 (Epping et al., 2011), was found in a consistent diagnosis-specific anchor and peak. Dysregulation of most of these genes was corroborated by independent studies. Using nine diagnosis and relapse pair samples, Nuno et al. found that *NCAM2*, *IFI44,* and *TSPYL5* significantly lost their accessibility upon relapse (Supplementary Figure S5). Finally, explored the relationship between the expression of these genes and clinical outcome, in terms of survival, using the Leucegene AML RNA-seq prognostic cohort (n = 373). We observed that high expression of these genes, including *TSPYL5*, negatively impacts the overall survival of AML patients, except for *CD70*, which did not show a significant effect (Supplementary Figure S6).

Discussion

AML aggressive subtypes are often linked to rearrangements in the mixed lineage leukemia gene (KMT2Ar). 10% of adult acute leukemias with a very poor prognosis and chemoresistance are caused by clinically significant and genetically well-defined

KMT2Ar (Krivtsov and Armstrong, 2007; Liedtke and Cleary, 2009; Meyer et al., 2018). Therefore, it is important to understand what genetic and epigenetic changes characterize relapse. Patients with t(9; 11)(p22; q23), the most frequent translocation which leads to the *KMT2A::MLLT3* fusion gene, carried by the patients in this study, show relatively acceptable results with intensive chemotherapy (Grimwade et al., 2010; Mrózek et al., 1997; Stölzel et al., 2016; Chen et al., 2013; Pigneux et al., 2015), placing them in the intermediate risk group according to ELN 2017 and ELN 2022 classifications (Döhner et al., 2017; 2022). This highlights the importance of identifying the oncogenic translocation for clinical decision making. Our chromatin conformation data revealed an undetected 9;11 translocation in one patient, indicating the need for more in-depth karyotyping using next-generation sequencing-based techniques. One option is targeted RNA-seq, which is becoming routine in diagnostics to detect low level fusion genes (Kerbs et al., 2022).

Advances in pharmacological inhibitors and targeted immunotherapies have considerably improved the treatment options for KMT2Ar leukemias (Issa et al., 2023; van der Sluis et al., 2023). However, KMT2Ar leukemias show highly heterogenous response to therapeutic regimens despite their similar oncogenic lesions (Issa et al., 2023; Lambo et al., 2023). Common genetic mechanisms leading to therapeutic resistance include clonal selection and the acquisition of secondary mutations. Additionally, a subset of these leukemias may also evade targeted therapies through epigenetic mechanisms that remain poorly understood (Tirtakusuma et al., 2022).

In this study, we explored epigenetic and transcriptomic changes in AML progression by integrating Omni-C, ATAC-seq, and RNA-seq data in two pairs of diagnosis and relapse samples. We found substantial chromatin remodeling, which was indicated by differences in the 3D chromatin structure, with a significant loss of chromatin interactions and open chromatin regions in relapse. Furthermore, differential gene expression analysis of those longitudinal samples revealed that upregulated genes are enriched in broad immune response pathways, including those related to cytokine signaling, especially interferon-alpha/beta signaling, while downregulated genes show enrichment in signaling pathways associated with Jak-STAT and interleukin.

Combined multi-omics data pinpointed three upregulated genes (*SDC2, NCAM2,* and *IFI44*) that showed negative effects on survival in a larger leukemia cohort. *TSPYL5* was downregulated in both of the relapse samples we tested, and higher expression of this gene was also associated with worse prognosis in the Leucegene cohort. In conclusion, our integrated analysis highlights distinct genomic and epigenomic profiles in relapse compared to diagnosis. Although the small sample size limits our study, the genes highlighted here showed dysregulation and had a prognostic effect in independent patient cohorts warranting their further exploration in larger cohorts to assess their clinical relevance and the therapeutic implications of these observations for KMT2Ar AML.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/geo/, GSE267375 https://www.ncbi.nlm.nih.gov/geo/, GSE267376.

Ethics statement

The study was approved by the Hamad Bin Khalifa University Institutional Review Board (QBRI-IRB-2020-02-017). The study was conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

NA: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing–original draft, Writing–review and editing. IC: Methodology, Supervision, Writing–review and editing. WV: Data curation, Formal Analysis, Software, Writing–review and editing. CC: Conceptualization, Data curation, Supervision, Writing–review and editing. SD: Conceptualization, Data curation, Supervision, Writing–review and editing. BM: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by a grant from Qatar National Research Funds (QNRF) - National Priority Research Program (NPRP): NPRP13S-0116-200088.

Acknowledgments

We are grateful to Roberto Gambato and Mirija Svaldi from the Hematology Laboratory in Bolzano Hospital for facilitating the patients' samples collection. The HPC resources and services used in this work were provided by the research computing group at Texas A&M University in Qatar. Research computing is funded by the Qatar Foundation for Education, Science and Community Development.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2024.1442539/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Each bar plot represents the average read coverage of cis interactions across chromosomes. **(A, B)** show the average coverage at diagnosis **(C, D)** show the average coverage at relapse. **(C)** is the relapse sample of **(A)**, **(D)** is the relapse sample of **(B)**.

SUPPLEMENTARY FIGURE S2
(A) Schematics of identifying relapse-specific differential interactions and peaks and diagnosis-specific interactions and peaks. **(B-F)** Examples of relapse-specific **(B, C)** and diagnosis-specific **(D–F)** Omni-C interactions and ATAC-seq peaks of UBE2J1 and PARP8; and BMP8A, ST8SIA1, and HPSE2, respectively. Red track shows H3K27me3, blue shows H3K27ac, and green shows H3K4me3/1 in Common Myeloid Progenitor – CD34 positive cells. **(G)** Bar plots showing the enrichment of biological processes and pathways based on significantly dysregulated genes between relapse versus diagnosis.

SUPPLEMENTARY FIGURE S3
(A) Schematics portraying the interactions of promoter regions with distal poised enhancers marked by H3K27me3. **(B)** Examples of gene promoters interacting with silencers. **(C)** Bar plots showing the proportion of diagnosis-specific and random interactions where the distal element is marked by H3K27me3. Error bars indicate the standard deviation of the random proportions based on 10,000 Monte Carlo simulations. **(D)** Box plots representing downregulated genes that overlapped with relapse-specific interactions and peaks and their comparison with the expression level of 10 different random sets of downregulated genes.**(D)** Box plots representing downregulated genes that overlapped with diagnosis-specific interactions and peaks and their comparison with the expression level of 10 different random sets of downregulated genes.

SUPPLEMENTARY FIGURE S4
Box plots representing ATAC-seq data from Nuno et al. of 18 stable initial diagnosis (blue) and relapse (brown) AML samples. 13 genes, which were upregulated and lost their interaction with a silencer. Diagnosis and relapse sample pairs are connected by grey lines. The log2FC differences and adjusted p-values are taken from the differential accessibility analysis between relapse and diagnosis using gene score,(a metric to estimate the accessibility of chromatin regions associated with specific genes).

SUPPLEMENTARY FIGURE S5
Box plots representing ATAC-seq data from Nuno et al. of 18 stable initial diagnosis (blue) and relapse (brown) AML samples. These five genes were found consistent across all three analyses. Diagnosis and relapse sample pairs are connected by grey lines. The log2FC differences and adjusted p-values were taken from the differential accessibility analysis between relapse and diagnosis using gene score, a metric to estimate the accessibility of chromatin regions associated with specific genes.

SUPPLEMENTARY FIGURE S6
Kaplan-Meier plots showing the overall survival (OS) of patients from the Leucegene Acute Myeloid Leukemia (AML) cohort based on gene expression levels of five different genes that we found significant. Each displays the survival probability over time, stratified by gene expression levels (above or below median). Patients with above-median gene expression levels of **(A)** SDC2, **(B)** CD70, **(C)** NCAM2, **(D)** IFI44, and **(E)** TSPYL5 show significant differences in survival compared to those with below-median levels, with p-values of 0.0004, 0.0194, 0.0001, 0.0135, and 0.0003, respectively.

SUPPLEMENTARY TABLE S1
Clinical characteristics of AML patients at diagnosis and relapse.

SUPPLEMENTARY TABLE S2
List of upregulated genes of relapse versus diagnosis AML samples with logFC >1 and adj-pval < 0.1.

SUPPLEMENTARY TABLE S3
List of downregulated genes of relapse versus diagnosis AML samples with logFC < −1 and adj-pval < 0.1.

References

Ahn, J.-S., Kim, H.-J., Kim, Y.-K., Lee, S.-S., Ahn, S.-Y., Jung, S.-H., et al. (2018). Assessment of a new genomic classification system in acute myeloid leukemia with a normal karyotype. *Oncotarget* 9 (4), 4961–4968. doi:10.18632/oncotarget.23575

Akl, M. R., Nagpal, P., Ayoub, N. M., Prabhu, S. A., Gliksman, M., Tai, B., et al. (2015). Molecular and clinical profiles of syndecan-1 in solid and hematological cancer for prognosis and precision medicine. *Oncotarget* 6 (30), 28693–28715. doi:10.18632/oncotarget.4981

Andrews, S., Biggins, L., Inglesfield, S., Carr, H., and Montgomery, J. (2010). FastQC A quality control tool for high throughput sequence data. Available at: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Bejanyan, N., Weisdorf, D. J., Logan, B. R., Wang, H.-L., Devine, S. M., de Lima, M., et al. (2015). Survival of patients with acute myeloid leukemia relapsing after allogeneic hematopoietic cell transplantation: a center for international blood and marrow transplant research study. *Biol. Blood Marrow Transplant. J. Am. Soc. Blood Marrow Transplant.* 21 (3), 454–459. doi:10.1016/j.bbmt.2014.11.007

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

Bottomly, D., Long, N., Schultz, A. R., Kurtz, S. E., Tognon, C. E., Johnson, K., et al. (2022). Integrative analysis of drug response and clinical outcome in acute myeloid leukemia. *Cancer Cell* 40 (8), 850–864.e9. doi:10.1016/j.ccell.2022.07.002

Canarte, V., Moser-Katz, T., Gavile, C., Barwick, B. G., Lee, K. P., and Boise, L. H. (2023). SDC2 expression is increased in myeloma cells in response to loss of pro-survival surface proteins, CD28 and CD86. *Blood* 142 (Suppl. 1), 3299. doi:10.1182/blood-2023-190453

Corces, M. R., Trevino, A. E., Hamilton, E. G., Greenside, P. G., Sinnott-Armstrong, N. A., Vesuna, S., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* 14 (10), 959–962. doi:10.1038/nmeth.4396

DiNardo, C. D., and Cortes, J. E. (2016). Mutations in AML: prognostic and therapeutic implications. *Hematol. Am. Soc. Hematol. Educ. Program* 2016 (1), 348–355. doi:10.1182/asheducation-2016.1.348

Döhner, H., Estey, E., Grimwade, D., Amadori, S., Appelbaum, F. R., Büchner, T., et al. (2017). Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood* 129 (4), 424–447. doi:10.1182/blood-2016-08-733196

Döhner, H., Wei, A. H., Appelbaum, F. R., Craddock, C., DiNardo, C. D., Dombret, H., et al. (2022). Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. *Blood* 140 (12), 1345–1377. doi:10.1182/blood.2022016867

Döhner, H., Weisdorf, D. J., and Bloomfield, C. D. (2015). Acute myeloid leukemia. *N. Engl. J. Med.* 373 (12), 1136–1152. doi:10.1056/NEJMra1406184

Durand, N. C., Shamim, M. S., Machol, I., Rao, S. S. P., Huntley, M. H., Lander, E. S., et al. (2016). Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* 3 (1), 95–98. doi:10.1016/j.cels.2016.07.002

Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., et al. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinforma. Oxf. Engl.* 21 (16), 3439–3440. doi:10.1093/bioinformatics/bti525

Epping, M. T., Meijer, L. A. T., Krijgsman, O., Bos, J. L., Pandolfi, P. P., and Bernards, R. (2011). TSPYL5 suppresses p53 levels and function by physical interaction with USP7. *Nat. Cell Biol.* 13 (1), 102–108. Article 1. doi:10.1038/ncb2142

Farrar, J. E., Schuback, H. L., Ries, R. E., Wai, D., Hampton, O. A., Trevino, L. R., et al. (2016). Genomic profiling of pediatric acute myeloid leukemia reveals a changing

mutational landscape from disease diagnosis to relapse. *Cancer Res.* 76 (8), 2197–2205. doi:10.1158/0008-5472.CAN-15-1015

Gröschel, S., Sanders, M. A., Hoogenboezem, R., Zeilemaker, A., Havermans, M., Erpelinck, C., et al. (2015). Mutational spectrum of myeloid malignancies with inv(3)/t(3;3) reveals a predominant involvement of RAS/RTK signaling pathways. *Blood* 125 (1), 133–139. doi:10.1182/blood-2014-07-591461

Howlader, N., Noone, A. M., Krapcho, M., Miller, D., Brest, A., Yu, M., et al. (2021). *SEER cancer statistics review, 1975-2018*. Bethesda, MD: National Cancer Institute. Available at: https://seer.cancer.gov/csr/1975_2018/index.html.

Issa, G. C., Aldoss, I., DiPersio, J., Cuglievan, B., Stone, R., Arellano, M., et al. (2023). The menin inhibitor revumenib in KMT2A-rearranged or NPM1-mutant leukaemia. *Nature* 615 (7954), 920–924. 7954. doi:10.1038/s41586-023-05812-3

Jan, M., Snyder, T. M., Corces-Zimmerman, M. R., Vyas, P., Weissman, I. L., Quake, S. R., et al. (2012). Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. *Sci. Transl. Med.* 4 (149), 149ra118. doi:10.1126/scitranslmed.3004315

Karlsson, L., Forestier, E., Hasle, H., Jahnukainen, K., Jónsson, Ó. G., Lausen, B., et al. (2017). Outcome after intensive reinduction therapy and allogeneic stem cell transplant in paediatric relapsed acute myeloid leukaemia. *Br. J. Haematol.* 178 (4), 592–602. doi:10.1111/bjh.14720

Kerbs, P., Vosberg, S., Krebs, S., Graf, A., Blum, H., Swoboda, A., et al. (2022). Fusion gene detection by RNA-sequencing complements diagnostics of acute myeloid leukemia and identifies recurring *NRIP1-MIR99AHG* rearrangements. *Haematologica* 107 (1), 100–111. Article 1. doi:10.3324/haematol.2021.278436

Kern, W., Haferlach, T., Schnittger, S., Ludwig, W. D., Hiddemann, W., and Schoch, C. (2002). Karyotype instability between diagnosis and relapse in 117 patients with acute myeloid leukemia: implications for resistance against therapy. *Leukemia* 16 (10), 2084–2091. doi:10.1038/sj.leu.2402654

Kishtagari, A., and Levine, R. L. (2021). The role of somatic mutations in acute myeloid leukemia pathogenesis. *Cold Spring Harb. Perspect. Med.* 11 (4), a034975. doi:10.1101/cshperspect.a034975

Krivtsov, A. V., and Armstrong, S. A. (2007). MLL translocations, histone modifications and leukaemia stem-cell development. *Nat. Rev. Cancer* 7 (11), 823–833. Article 11. doi:10.1038/nrc2253

Lambo, S., Trinh, D. L., Ries, R. E., Jin, D., Setiadi, A., Ng, M., et al. (2023). A longitudinal single-cell atlas of treatment response in pediatric AML. *Cancer Cell* 41 (12), 2117–2135.e12. doi:10.1016/j.ccell.2023.10.008

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., et al. (2013). Software for computing and annotating genomic ranges. *PLOS Comput. Biol.* 9 (8), e1003118. doi:10.1371/journal.pcbi.1003118

Levin, M., Stark, M., Ofran, Y., and Assaraf, Y. G. (2021). Deciphering molecular mechanisms underlying chemoresistance in relapsed AML patients: towards precision medicine overcoming drug resistance. *Cancer Cell Int.* 21 (1), 53. doi:10.1186/s12935-021-01746-w

Li, D., Purushotham, D., Harrison, J. K., Hsu, S., Zhuo, X., Fan, C., et al. (2022). WashU Epigenome browser update 2022. *Nucleic Acids Res.* 50 (W1), W774–W781. doi:10.1093/nar/gkac238

Liedtke, M., and Cleary, M. L. (2009). Therapeutic targeting of MLL. *Blood* 113 (24), 6061–6068. doi:10.1182/blood-2008-12-197061

Lu, L., Liu, X., Huang, W.-K., Giusti-Rodríguez, P., Cui, J., Zhang, S., et al. (2020). Robust Hi-C maps of enhancer-promoter interactions reveal the function of non-coding genome in neural development and diseases. *Mol. Cell* 79 (3), 521–534. doi:10.1016/j.molcel.2020.06.007

Masetti, R., Castelli, I., Astolfi, A., Bertuccio, S. N., Indio, V., Togni, M., et al. (2016). Genomic complexity and dynamics of clonal evolution in childhood acute myeloid leukemia studied with whole-exome sequencing. *Oncotarget* 7 (35), 56746–56757. doi:10.18632/oncotarget.10778

Metzeler, K. H., Herold, T., Rothenberg-Thurley, M., Amler, S., Sauerland, M. C., Görlich, D., et al. (2016). Spectrum and prognostic relevance of driver gene mutations in acute myeloid leukemia. *Blood* 128 (5), 686–698. doi:10.1182/blood-2016-01-693879

Meyer, C., Burmeister, T., Gröger, D., Tsaur, G., Fechina, L., Renneville, A., et al. (2018). The MLL recombinome of acute leukemias in 2017. *Leukemia* 32 (2), 273–284. Article 2. doi:10.1038/leu.2017.213

Northcott, P. A., Lee, C., Zichner, T., Stütz, A. M., Erkek, S., Kawauchi, D., et al. (2014). Enhancer hijacking activates GFI1 family oncogenes in medulloblastoma. *Nature* 511 (7510), 428–434. doi:10.1038/nature13379

Nuno, K., Azizi, A., Koehnke, T., Lareau, C., Ediriwickrema, A., Corces, M. R., et al. (2024). Convergent epigenetic evolution drives relapse in acute myeloid leukemia. *eLife* 13, e93019. doi:10.7554/eLife.93019

Papaemmanuil, E., Gerstung, M., Bullinger, L., Gaidzik, V. I., Paschka, P., Roberts, N. D., et al. (2016). Genomic classification and prognosis in acute myeloid leukemia. *N. Engl. J. Med.* 374 (23), 2209–2221. doi:10.1056/NEJMoa1516192

Rasche, M., Zimmermann, M., Steidel, E., Alonzo, T., Aplenc, R., Bourquin, J.-P., et al. (2021). Survival following relapse in children with acute myeloid leukemia: a report from AML-BFM and cog. *Cancers* 13 (10), 2336. Article 10. doi:10.3390/cancers13102336

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007

Robinson, J. T., Turner, D., Durand, N. C., Thorvaldsdóttir, H., Mesirov, J. P., and Aiden, E. L. (2018). Juicebox.js provides a cloud-based visualization system for Hi-C data. *Cell Syst.* 6 (2), 256–258. doi:10.1016/j.cels.2018.01.001

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26 (1), 139–140. doi:10.1093/bioinformatics/btp616

Schlenk, R. F., Jaramillo, S., and Müller-Tidow, C. (2018). Improving consolidation therapy in acute myeloid leukemia—a tough nut to crack. *Haematologica* 103 (10), 1579–1581. doi:10.3324/haematol.2018.200485

Schwenger, E., and Steidl, U. (2021). An evolutionary approach to clonally complex hematologic disorders. *Blood Cancer Discov.* 2 (3), 201–215. doi:10.1158/2643-3230.BCD-20-0219

Siegel, R. L., Miller, K. D., Wagle, N. S., and Jemal, A. (2023). Cancer statistics, 2023. *CA A Cancer J. Clin.* 73 (1), 17–48. doi:10.3322/caac.21763

Stein, E. M., DiNardo, C. D., Pollyea, D. A., Fathi, A. T., Roboz, G. J., Altman, J. K., et al. (2017). Enasidenib in mutant IDH2 relapsed or refractory acute myeloid leukemia. *Blood* 130 (6), 722–731. doi:10.1182/blood-2017-04-779405

Stone, R. M., Mandrekar, S. J., Sanford, B. L., Laumann, K., Geyer, S., Bloomfield, C. D., et al. (2017). Midostaurin plus chemotherapy for acute myeloid leukemia with a FLT3 mutation. *N. Engl. J. Med.* 377 (5), 454–464. doi:10.1056/NEJMoa1614359

Tirtakusuma, R., Szoltysek, K., Milne, P., Grinev, V. V., Ptasinska, A., Chin, P. S., et al. (2022). Epigenetic regulator genes direct lineage switching in MLL/AF4 leukemia. *Blood* 140 (17), 1875–1890. doi:10.1182/blood.2021015036

van der Sluis, I. M., de Lorenzo, P., Kotecha, R. S., Attarbaschi, A., Escherich, G., Nysom, K., et al. (2023). Blinatumomab added to chemotherapy in infant lymphoblastic leukemia. *N. Engl. J. Med.* 388 (17), 1572–1581. doi:10.1056/NEJMoa2214171

Verma, D., Kantarjian, H., Faderl, S., O'Brien, S., Pierce, S., Vu, K., et al. (2010). Late relapses in acute myeloid leukemia: analysis of characteristics and outcome. *Leukemia and Lymphoma* 51 (5), 778–782. doi:10.3109/10428191003661852

Vicente-Dueñas, C., Hauer, J., Cobaleda, C., Borkhardt, A., and Sánchez-García, I. (2018). Epigenetic priming in cancer initiation. *Trends Cancer* 4 (6), 408–417. doi:10.1016/j.trecan.2018.04.007

Wang, B., Liu, Y., Hou, G., Wang, L., Lv, N., Xu, Y., et al. (2016). Mutational spectrum and risk stratification of intermediate-risk acute myeloid leukemia patients based on next-generation sequencing. *Oncotarget* 7 (22), 32065–32078. doi:10.18632/oncotarget.7028

Zhang, S., Plummer, D., Lu, L., Cui, J., Xu, W., Wang, M., et al. (2022). DeepLoop robustly maps chromatin interactions from sparse allele-resolved or single-cell Hi-C data at kilobase resolution. *Nat. Genet.* 54 (7), 1013–1025. Article 7. doi:10.1038/s41588-022-01116-w

Zhu, H., Uusküla-Reimand, L., Isaev, K., Wadi, L., Alizada, A., Shuai, S., et al. (2020). Candidate cancer driver mutations in distal regulatory elements and long-range chromatin interaction networks. *Mol. Cell* 77 (6), 1307–1321. doi:10.1016/j.molcel.2019.12.027