Cisplatin reprogramming of protein phosphorylation

Luyu Qi[1,2,#], Yang Zhou[1,2,#], Yinzhu Hou[1,2], Qun Luo[1,2], Yan Xu[2], Xingkai Liu[1], Wanchen Yu[3], Bobo Xin,[4] Yaolong, Huang[4], Juanjuan Han[5], Xiangjun Li[2], Yanyan Zhang[1], Feifei Jia[1], Tiantian Fang[1], Shijun Wang[3,]*, Peter J. Sadler[6,]*, Yao Zhao[1,2]*, Fuyi Wang[1,2,3,5]*

[1] Beijing National Laboratory for Molecular Sciences; CAS Research/Education Center for Excellence in Molecular Sciences; National Centre for Mass Spectrometry in Beijing; CAS Key Laboratory of Analytical Chemistry for Living Biosystems, Institute of Chemistry, Chinese Academy of Sciences, Beijing, 100190, People's Republic of China
[2] University of Chinese Academy of Sciences, Beijing 100049, People's Republic of China
[3] Basic Medical College, Shandong University of Traditional Chinese Medicine, Jinan 250355, P. R. China
[4] NGS Department, Sangon Biotech (Shanghai), Shanghai 201611, P. R, China
[5.] National Centre for Mass Spectrometry in Beijing, 100190, People's Republic of China
[6] Department of Chemistry, University of Warwick, Coventry CV4 7AL, United Kingdom
[#]These authors contributed equally to this work.
*Correspondence: fuyi.wang@iccas.ac.cn (F.Y.W.), yaozhao@iccas.ac.cn (Y.Z.), p.j.sadler@warwick.ac.uk (P.J.S.), wsj@sdutcm.edu.cn (S.J.W.)

ABSTRACT

Cisplatin is a DNA-targeting chemotherapeutic. We have mapped cisplatin damage to specific genes in human lung cancer cells. Surprisingly, cisplatin targets the majority of protein kinase genes and protein phosphatase genes in the human genome. This suggests that cisplatin can reprogram protein phosphorylation genome-wide. We have profiled the protein expression and phosphorylation in human testicular cancer cells subjected to cisplatin treatment, since testicular cancer is curable by cisplatin. We demonstrate that most downregulated proteins are encoded by cisplatin damaged genes. These proteins include a series of protein kinases and protein phosphatases, leading to significant changes in the phosphorylation level of >600 proteins in testicular cancer cells. Hence, reprograming of protein phosphorylation is proteome-wide. Importantly, the reprogrammed protein phosphorylation activates G2/M DNA-damage checkpoint regulation and ATM signalling pathways, causing cell cycle arrest at the G2/M phase. These findings suggest that specific protein phosphorylation pathways are potential new targets for platinum drug design.

INTRODUCTION

Cisplatin, *cis*-[$PtCl_2(NH_3)_2$], is a first-line chemotherapeutic drug, in particular by systemic administration in combination therapy for solid tumours[1,2]. After hydrolysis in cells, cisplatin binds strongly to a variety of biological molecules, e.g., proteins, thiol-containing small molecules, and nucleic acids[3,4]. The major target is believed to be nuclear DNA (nDNA)[3,5], although the formation of nDNA adducts is not the sole determinant of cisplatin-induced cytotoxicity[6] and has been linked to side effects and toxicity[7]. The formation of 1,2-intrastrand -GG-/-AG- (90%) adducts, 1,3-intrastrand -GNG- adducts and interstrand crosslinks unwinds and bends DNA helices, inhibiting DNA replication and transcription, and in turn inducing cell apoptosis and death[3].

The human high mobility group box 1 (HMGB1) protein is a non-specific DNA binding protein, and well recognized for its specific binding to cisplatin-crosslinked double–stranded DNA in both cell-free media[8,9] and in cells[9-11]. HMGB1 functions as a protector of cisplatin-induced DNA damage from nucleotide excision repair (NER)[5], and may serve as a negative transcription cofactor to block the binding of transcription factors such as Smad3 to DNA, as revealed by single cell imaging[11]. Box a domain of HMGB1 has been demonstrated to have a higher affinity than box b for 1,2-cisplatin-crosslinked DNA[9,12]. Based on such a high affinity of the HMGB1a domain for cisplatin-bound DNA, Shu et al. developed a "cisplatin-seq" method to map cisplatination sites with a single base resolution throughout the whole genome[12]. They revealed that cisplatin was more likely to attack mitochondrial DNA (mDNA) than nDNA, and that the cisplatin damage was enriched within promotors and regions harbouring transcription termination sites[12].

Despite technical challenges arising from the extremely low incidence of cisplatin damage on DNA on a genome-wide scale[13], and impacts of cisplatin damage on the PCR-based sequencing, e.g., stalling PCR process and making base mismatch, there have been a number of innovative strategies that provide insights into the cisplatin damage sites at the genome-wide level[14]. Apart from the cisplatin-seq approach described above, Hu et al. developed a set of methods, termed "damage-seq" and "eXcision Repair-seq" (XR-seq), to map, at single-nucleotide resolution, the location and repair of cisplatin lesions in genomic DNA adducts[15], respectively. Their mapping data demonstrated that cisplatin damage was uniform and dominated by the underlying sequence, but the repair of cisplatin lesions was heterogeneous throughout the whole genome and affected by multiple factors such as transcription and chromatin states. Therefore, they thought that the overall effect of cisplatin damage in the genome is not primarily determined by cisplatin damage, but by the effectiveness of DNA damage repair[15]. However, similar to the cisplatin-seq approach, the damage-seq method was not able to assign cisplatin lesions to specific gene loci. To address this issue, the same group integrated their "damage-seq", "XR-seq" and RNA-seq approaches to study comprehensively the genome-wide profiles of cisplatin lesions in mouse organs[16]. They revealed that the expression of 12 genes was commonly downregulated due to cisplatin damage across mouse kidney, liver, lung, and spleen[16]. However, the combined use of transcriptomics strategies did not produce a full map of cisplatin lesions throughout the whole genome due to the tissue specificity of transcriptomics.

The aim of the research reported here was to map the exact genomic distribution of cisplatin-damaged gene loci by using forward chemical genetics strategy[17,18]. The biotin-streptavidin system is often applied to enrich and isolate target DNA and proteins at ultra trace levels in complicated biological samples[17-21]. However, to avoid the effect of biotin modification on the chemical properties and pharmacological functions of the drug cisplatin itself, we combined HMGB1 box a (HMGB1a)-based affinity microprobes and high throughput sequencing to screen gene-lesion loci induced by cisplatin on a whole-genome scale (Fig. 1), and to correlate the gene lesions with the action of the drug. Significantly, the data suggest that cisplatin may reprogram protein phosphorylation in cells may have a major influence on its pharmacological function. We further carried out global proteomics / phosphoproteomics analysis of NCCIT human testicular cancer cells subjected to cisplatin treatment to explore the cisplatin induced reprogramming of protein phosphorylation and its role in the mechanisms of action this first line anticancer drug. These findings suggest that specific protein phosphorylation pathways may become new targets for platinum drug design with the aim of increasing efficacy and reducing side effects.

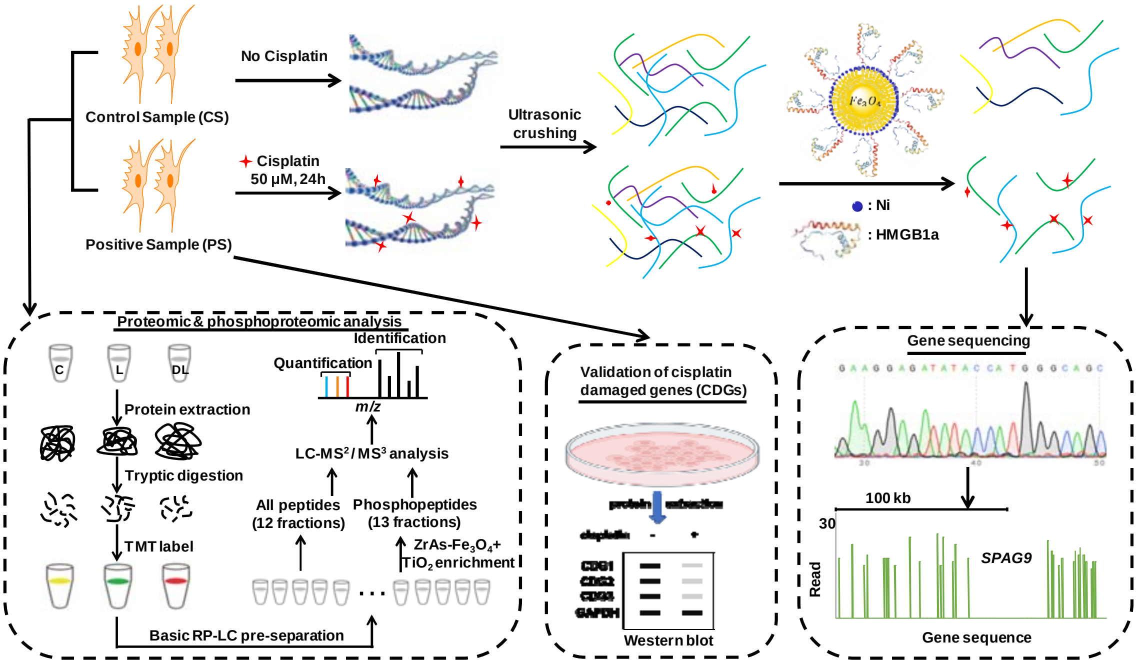
Control Sample (CS)
No Cisplatin
Positive Sample (PS)
Cisplatin
50 μM, 24h
Ultrasonic crushing
Fe3O4
: Ni
: HMGB1a
Proteomic & phosphoproteomic analysis
C
L
DL
Protein extraction
Tryptic digestion
TMT label
Basic RP-LC pre-separation
Identification
Quantification
m/z
LC-MS2/MS3 analysis
All peptides (12 fractions)
Phosphopeptides (13 fractions)
ZrAs-Fe3O4+ TiO2 enrichment
Validation of cisplatin damaged genes (CDGs)
protein extraction
cisplatin
-
+
CDG1
CDG2
CDG3
GAPDH
Western blot
Gene sequencing
G A A G G A G A T A T A C C A T G G G C A G C
30
40
50
100 kb
30
Read
SPAG9
Gene sequence

Fig. 1. Diagrammatic illustration of the workflow for capture and mapping of gene loci attacked by cisplatin. The procedure includes cell culture, extraction and fragmentation of genomic DNA, capture, sequencing and alignment of cisplatin-damaged DNA fragments, identification of cisplatin-damaged genes (CDGs), validation of CDGs by Western Blot assays, and MS-based quantitative proteomics and phosphoproteomics analysis.

RESULTS.

Enrichment and sequencing of cisplatin-damaged DNA

We functionalized Ni-based magnetic microbeads with HMGB1a to assemble affinity microprobes for capturing and isolating 1,2-cisplatin-crosslinked DNA[3,9] from genomic DNA fragments derived from human A549 non-small cell lung cancer cells treated with 50 μM cisplatin for high throughput Next Generation Sequencing (NGS) (Fig. 1). Details of the expression and purification of HMGB1a, construction of affinity microprobes, cell harvesting, and DNA extraction and fragmentation are provided in Methods and Details.

Since HMGB1 binds to cisplatin-crosslinked DNA with high affinity, but can also bind to bent native DNA[3], we extracted and fragmented genomic DNA from A549 cells with (positive samples) and without (control samples) exposure to cisplatin for DNA sequencing. ICP-MS analysis for Pt showed that the platination level of genomic DNA extracted from cisplatin-treated cells reached 320 Pt atoms per 10^6 base pairs (Fig. 2A), consistent with our previous report[13]. Furthermore, gel-electrophoresis data demonstrated that the size of the DNA fragments gained from both the positive sample (PS) and the control sample (CS) ranged from 250 to 1000 bp, but the number of larger DNA fragments from positive samples was less than those from control samples (Fig. 2B). This is likely ascribed to platination, which could render genomic DNA more susceptible to sonication and more prone to aggregation[22]. Given that HMGB1a binds to both non-platinated and platinated DNA, it is not surprising that the same amount of the microprobes captured similar amounts of DNA fragments from positive and control samples (Extended Data Table 1). However, the subsequent sequencing revealed that there were significant differences between the DNA fragments enriched from the positive and control samples.

Next, we constructed DNA libraries based on the captured DNA fragments. Considering that cisplatin damage on DNA may stall DNA replication and/or cause base mismatch[14,23]. In this work, we tried to remove Pt bound to DNA bases by using excess of GSH (glutathione) because Pt(II) has a high affinity for S-donor ligands[24]. However, we did not observe significant difference between the amplification rates of GSH-treated and intact platinated DNA (data not shown). Moreover, previous reports also demonstrated that the replication of DNA can still reach a similar level for cisplatin-damaged DNA compared to intact DNA by extension of the replication time[23]. Therefore, we extended the PCR expansion time of both positive and negative samples to 30 min. The results showed that the DNA sequences in DNA libraries of the positive samples and control samples have similar sizes (Fig. 2C). This means that possible Pt-modification-induced bias for PCR amplification of DNA was circumvented (details given in Methods and Details), which minimized the risk of false negative results from gene mapping. Indeed, the total read counts and total base counts of the positive samples did not show a pronounced difference from those of the control samples. However, because cisplatin prefers to crosslink -GG- sites on DNA, the GC base-pair counts of the positive samples accounted for a slightly higher ratio to the total base counts than those of the control samples (Fig. 2D; Extended Data Table 1). Consequently, the mapping rates of total reads in proper base pairs of positive samples were slightly lower than those of control samples (Fig. 2E; Extended Data Table 1). Nevertheless, there are no differences in the total mapping rates (≈99%) and unique mapping rates (>91%) of genomic DNA fragments from both positive and control samples (Fig. 2E; Extended Data Table 1).

To identify cisplatin-damaged genes (CDGs), we compared the total reads of all peaks mapped to a gene in the positive sample (PS) with those of all peaks mapped to the same gene in the corresponding control sample (CS). If the ratio ($[\text{number of peaks}]_{PS}/[\text{number of peaks}]_{CS}$), the fold-enrichment ($FE_P$) of a peak mapped to a gene, was ≥ 1.5, the peak was counted to the gene. The fold-enrichment of a gene (FE_G) is the sum of the FE_P of all peaks mapped to the gene (details in Methods and Details). Notably, to eliminate the bias arising from the different total reads of all peaks mapped in all samples, we used the normalized value of the read of a peak to the total read of all peaks for the sample (Extended data Table 1). Following this method, we mapped 246,424 and 362,594 peaks with $FE_P \geq 1.5$ from replicates 1 and 2, respectively. As shown in Fig. 2F, the peaks mapped to promoter regions, located 3k base pairs upstream of a gene, account for the highest percentage (32.48%) of all peaks mapped genome-wide. Considering the different lengths of each region in the genome, the coverage of peaks mapped to the promotor regions of all CDGs account for 82.5% of the total length of promotor region 1 (≤ 1 kb), 35.0% of promotor region 2 (1-2 kb), 27.8% of promotor region 3 (2-3 kb) (Fig. 2G). These coverages are much higher than for peaks mapped to other regions, e.g. exon and intron regions. These results show that cisplatin damage is enriched in the promotor regions, in particular, in regions close to encoding regions, in line with previous reports[12]. Since HMGB1a binds to double-stranded DNA over 5 base-pairs[3], we calculated the occurrence frequency of pentanucleotide motifs for all peaks containing platination sites, and found that the top 5 most-frequently platinated pentanucleotide motifs are CTGGG, CCAGG, GCTGG, GGAGG and CAGGC (Fig. 2H).

When we calculated the FE of a CDG, we discarded the peaks mapped to the distal intergenic region of a gene. Following this procedure, we mapped a total of 16216 CDGs with $FE_G \geq 1.5$ which were common in the two replicates (Fig. 2I; Extended Data Table 2). Moreover, the fold-enrichments of CDGs mapped in the two replicates were linearly correlated, with R = 0.804, as shown in Fig. 2J, evidencing good reproducibility of our sequencing. The four CDGs with the highest FE_G mapped in this way are shown in Extended Data Fig. 1. Remarkably, *PTPRN2,* the CDG with an FE_G of 33.3 (Extended Data Table 2), contains 82 (A/T)GGGG repeats instead of the GG/GGG motifs (Fig. 2H). This is consistent with the well-established finding that cisplatin prefers to bind to GG sites on DNA[3,5].

The 16216 CDGs are unevenly distributed over all 23 pairs of chromosomes (Fig. 2K), showing a moderate dependence on the average G-density of the chromosomes, i.e., the higher the average G-density of a chromosome, the more the genes are attacked by cisplatin (Fig. 2L), again consistent with the known preference of cisplatin for crosslinking -GG- sites. It is notable that due to the large difference in the length of genes, the fold-enrichment (FE_G) of a CDG based on the fold-enrichment of all peaks to the gene does not correspond simply to the base coverage of all identified peaks for the gene (Extended Data Table 2). For example, 198 peaks were identified for gene named *TRAPPC9,* which has the highest FE_G (476.4) but with a low base coverage of 15.5%. In contrast, for microRNA gene *MIR182* with a FE_G of 7, only 3 peaks were identified, but the base coverage of these peaks to the whole gene of *MIR182* was 74.7% (Extended Data Table 2). With regard to this, the FE_G in this work is only an indicator of CDG-mapping confidence, but not of the degree of cisplatin damage to a gene.

Ingenuity Pathway Analysis (IPA) database was employed to perform genome-wide annotation on the mapped CDGs. The 16216 CDGs were input into the IPA data pool and match to 11952 protein encoding genes in the human genome (hg19), including 2121 enzymes, 1220 transcription regulators, 723 transporters, 591 kinases and 197 phosphatases (Fig. 2M, Extended Data Table 2), as well as 11 microRNA (miRNA) encoding genes, of which the majority have relatively low fold-enrichment as their sequences are much shorter than genomic DNA. Notably, among the 591 cisplatin-damaged kinase genes, 400 are protein kinase genes (CDPKGs, Extended Data Table 3), accounting for 77% of 518 reported putative protein kinase genes in the human genome[25]. There are also 131 protein phosphatase genes (CDPPGs) among the 210 cisplatin-damaged phosphatase genes (Extended Data Table 4), accounting for 96% of the 136 reported protein phosphatases in the human genome. Significantly, among the CDGs, 418 genes belong to the 628 prior therapeutic gene targets for cancers identified by the CRISPR/Cas9-based screening, 54 (13%) of which are targets of clinically-used drugs or candidates in preclinical development[26] (Extended Data Table 5). Also, the average percentage of G-bases in CDPKGs and CDPPGs is significantly higher than that of all the genes in the human genome (hg19, Fig. 2N). Therefore, platination on the whole genome is preferential for genes with high G-density.

For HeLa cells, it has been shown that mitochondrial DNA (mDNA) is more susceptible to cisplatin attack than nDNA[6,12]. However, here for lung cancer cells, we detected only 14 cisplatin-damaged mitochondrial genes with low FE_G ranging from 1.5 to 2.4 (Extended Data Table 6). Considering the much lower G-density in mDNA (Fig. 2N), it is reasonable to assume that mitochondrial DNA has a low platination level. The differences between the previous and present work include the cell type (HeLa ovarian cancer cells[12] and immortalized lymphocytes[15], compared to A549 lung cancer cells in the present work), time of treatment (3 – 24 h[12] and 1.5 h[15], compared to 24 h here) and the dose of cisplatin (50 μM[12] and 200 μM[15], compared to 50 μM here), mapping methods (Cisplatin-seq[12] and Damage-seq[15], compared to NGS used here). It is notable that among the 16216 CDGs with FE_G >1.5, there are 36 pseudogenes (Extended Data Table 2). In general, pseudogenes are thought to be non-functional "junk genes". However, increasing evidence argues that pseudogenes can act as key regulators at DNA, RNA or protein levels in various human diseases such as cancer[27,28].

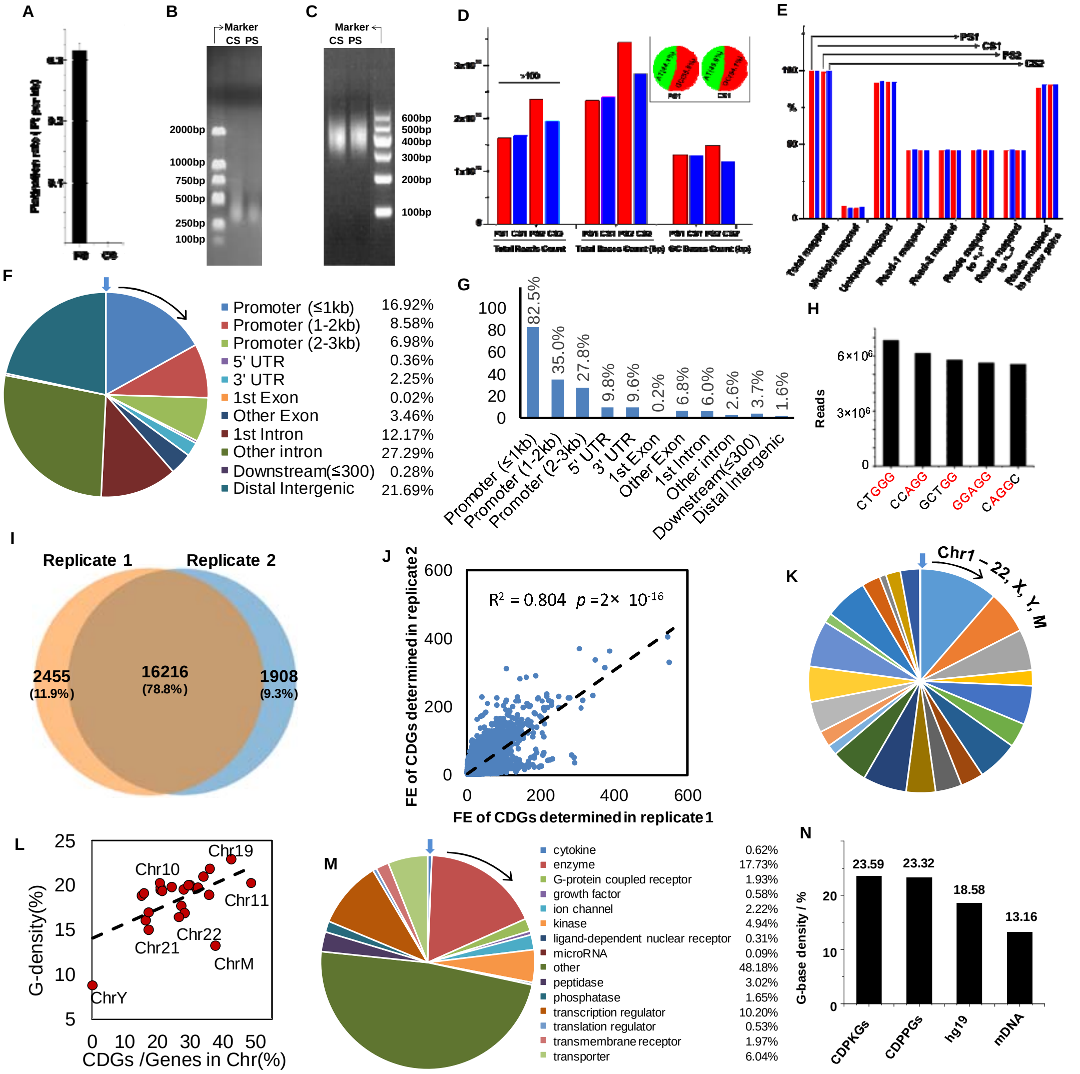

Fig. 2. Mapping of cisplatin damaged genes. **(A)** The platination level of DNA fragments extracted from positive sample (PS) and control sample (CS), determined by ICP-MS. **(B)** The size distribution of DNA fragments from positive sample (PS) and control sample (CS). **(C)** The size distribution of sequences in DNA libraries constructed from captured DNA fragments. **(D)** Total reads, total bases and GC bases mapped for positive and control samples. **(E)** Mapping rates counted for positive and control samples. **(F)** Distribution of cisplatin-damaged DNA fragments in genomic regions. **(G)** Ratio of the lengths of the cisplatin-damaged DNA fragments in specific genomic regions. **(H)** The five pentanucleotide motifs that were the most frequently detected in CDGs. **(I)** Venn Diagram of total cisplatin-damaged genes (CDGs) with FE_G >1.5 in two independent biological replicates. **(J)** Correlation analysis of the normalized fold-enrichment of CDGs in two replicates. **(K)** Distribution of CDGs over chromosomes. **(L)** Correlation of the relative G-density in the chromosome (y-axis) with the proportion of CDGs detected in a chromosome to total number of genes in human genome (x-axis). **(M)** Biological function classification of 11952 protein-encoding cisplatin-damaged genes (CDGs) with FE_G >1.5. **(N)** Average percent of G bases in CDPKGs and CDPPGs, hg19 and mDNA.

Western Blot assays for cisplatin treated cancer cells

To verify the genome-wide mapping results of CDGs, we performed classical Western Blot assays to measure the gene expression level of 15 protein kinase genes and *SPAG9* (sperm associated antigen 9) with or without cisplatin exposure (Extended Data Table 7). *SPAG9* abundantly expresses C-jun-amino-terminal kinase-interacting protein 4 (JIP-4) which consists of four subunits (Extended Data Fig. 2A) in testicular haploid germ cells and is essential for the development, migration and invasion of cancer[29]. JIP-4 as a scaffolding protein knits mitogen-activated protein kinases (MAPKs) and their transcription factors to activate specific phosphorylation signalling pathways, such as MAPK and PI3K pathways.[30]

As shown in Fig. 3A, the expression levels of 14 selected genes decreased after cisplatin treatment. The changes in expression levels of 8 genes, including 7 protein kinase genes and *SPAG9*, were significant ($P < 0.05$) or very significant ($P < 0.01$) (Fig. 3B). The data indicate that these genes were indeed damaged by cisplatin, substantiating the gene mapping results described above. However, there is no direct correlation between the fold-enrichment of a specific CDG, or the base-coverage of cisplatin-damaged peaks over a damaged gene and the fold-change in expression of the gene (Extended DATA Fig. 3 and Extended Data Table 7). This is most probably attributable to the complexity of the gene expression machinery. The expression level of a specific gene is greatly influenced by the location of cisplatin damage in the promoter or exon regions.

Indeed, when we further analysed the cisplatin-damaged sites on *SPAG9* (Extended Data Table 8), we found that (i) no cisplatin lesions were mapped on the promotor sequences; and (ii) cisplatin damaged the non-encoding regions (introns) much more than the encoding regions of the gene (exons, Extended Data Fig. 2B). These make it possible that although the full-length JIP-4 decreased in abundance, one or two subunits of JIP-4 increased in abundance (Fig. 3A).

Given that *SPAG9* specifically expresses JIP-4 in high abundance in testicular haploid germ cells, the cisplatin damage may account for its high cure rate in testicular cancer. Indeed, we found that siRNA silencing of the *SPAG9* gene (Fig. 3C) induced a similar level of early- and late-stage apoptosis in NCCIT cells subjected to cisplatin treatment (Fig. 3D, E and F). This demonstrated that *SPAG9* plays an important role in the action of cisplatin as the upstream mediator of phosphorylation pathways (Extended Data Fig. 4).

The expression of the receptor tyrosine kinase *MERTK* was the most strongly inhibited (FC = 32) due to cisplatin damage (Fig. 3A – B). This receptor phosphorylates Akt1-Y26, mediating Akt activation and survival signalling, which in turn drives oncogenesis and therapeutic resistance[31]. Thus, cisplatin damage to *MERTK* may be implicated in mechanism of action of cisplatin. Hence, *MERTK* is also a potential therapeutic target[31,32].

The expression of *MERTK*-related gene *AKT3* was inhibited 3.5-fold by cisplatin (Fig. 3A – B). *AKT3* encodes a member of the AKT serine/threonine protein kinase family. As a middle-stream regulator of IGF1 signalling, AKT3 plays a key role in signalling and regulating cell survival in insulin signalling, angiogenesis and tumour formation. Cisplatin-induced damage to this gene could trigger a series of downstream changes, inducing apoptosis and death of cancer cells[33].

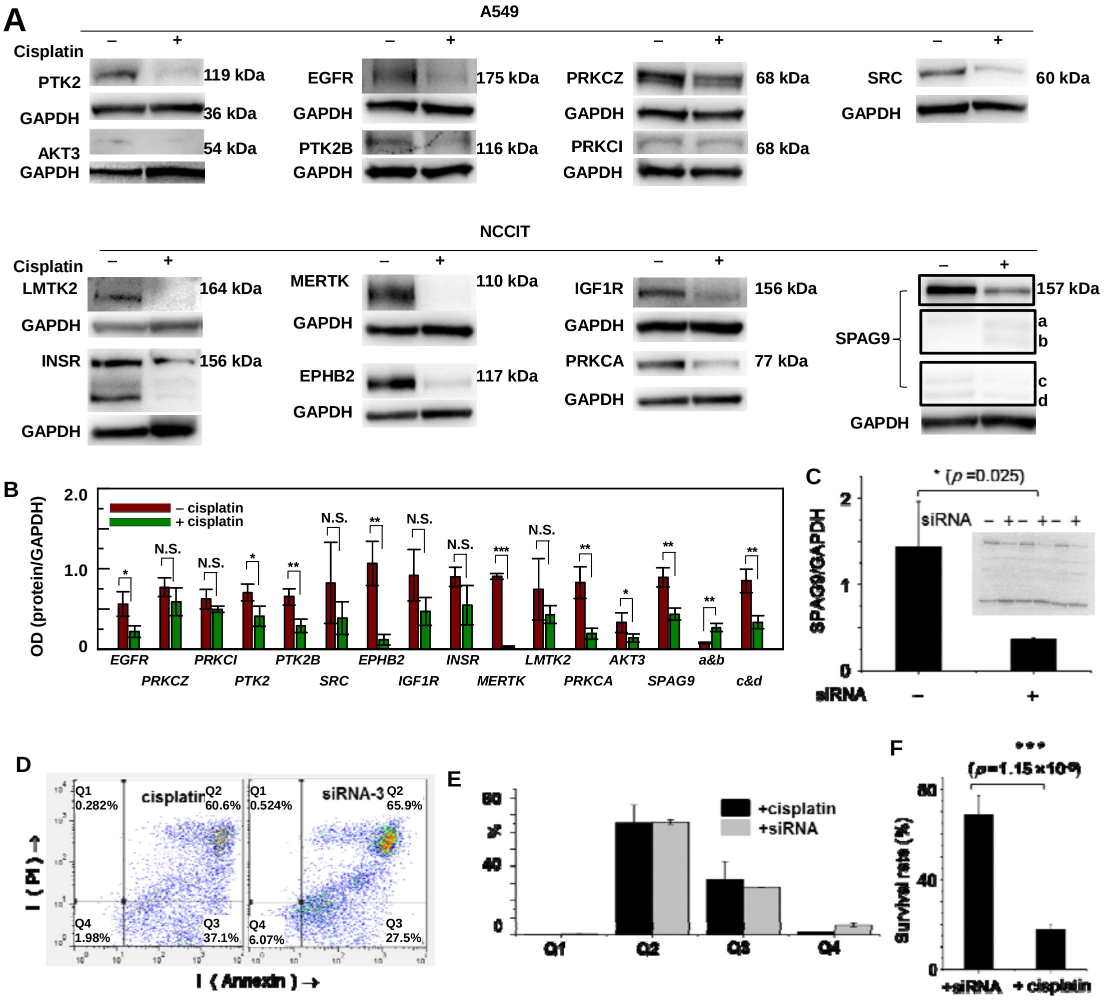

Fig. 3. Western Blot assays and siRNA silencing. (A) Western Blot images of selected gene-encoding proteins expressed in A549 lung and NCCIT testicular cancer cells with (+) and without (−) cisplatin (12 μM) treatment. GAPDH was the internal reference. The a – d subunits of JIP-4 (expressed by *SPAG9*) are shown in Extended Data Fig. 2A. (B) Optical density (OD) ratio of a gene-encoded protein to GAPDH. (C) Western Blot data for siRNA silencing of *SPAG9*. (D) Flow cytometric analysis of viability (Q1), early-stage apoptosis (Q2), late-stage apoptosis (Q3) and necrosis (Q4) for NCCIT cells subjected to cisplatin treatment in comparison with siRNA silencing of *SPAG9*. (E) The quantitative analysis of the flow cytometric analysis shown in D. (F) Survival rate (%) of NCCIT cells subjected to cisplatin treatment or siRNA silencing of *SPAG9*. In B, two-tailed unpaired Student's t-test was used to statistics process of the Western Blot assay (n = 3). N.S.: no significant; * indicates $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. $P_{EGFR}=0.0244$, $P_{PRKCZ}=0.2057$, $P_{PRKCI}=0.1615$, $P_{PTK2}=0.0366$, $P_{PTK2B}=0.0088$, $P_{SRC}=0.2343$, $P_{EPHB2}=0.0044$, $P_{IGF1R}=0.1000$, $P_{INSR}=0.0893$, $P_{MERTK}=2.71\times10^{-6}$, $P_{LMTK2}=0.2453$, $P_{PRKCA}=0.0062$, $P_{AKT3}=0.0126$, $P_{SPAG9\text{-full}}=0.0049$, $P_{SPAG9\text{-a\&b}}=0.0030$, $P_{SPAG9\text{-c\&d}}=0.0061$.

Proteomic analysis of testicular cancer cells

To address the impact of cisplatin damage on gene expression in a large scale, we performed a mass spectrometry (MS)-based global quantitative proteomics study on human testicular NCCIT cancer cells subjected to cisplatin treatment (Extended data table 9). Proteins were extracted from the living NCCIT cells (L group), and from the dead + living NCCIT cells (DL group), then identified and quantified by high resolution MS coupled to nanoLC after trypsin digestion, in comparison with the proteins extracted from the NCCIT cells without cisplatin treatment (control, C group). The details for proteomic analysis are given in the experimental section and Extended Data Fig. 5.

We identified 4615 proteins which were expressed in all three biological replicates of each group (Fig. 4A). In the DL group for the cisplatin-treated cells, compared to the C group, 350 proteins were downregulated and 253 proteins were upregulated (Fig. 4B, Extended DATA Fig. 6). In the L group, 184 proteins were downregulated and 222 proteins were upregulated upon cisplatin treatment (Fig. 4C) with a fold change (FC) of > 1.2. If the abundance ratio (AR) of a protein expressed in the cisplatin-treated groups to the control group is > 1, the FC is equal to AR. When AR < 1, the FC is calculated as the negative reciprocal of AR (−1/AR). As predicted, more proteins were up- or down-regulated in the DL group than in the L group (Fig. 4D), indicating that the expression of more proteins in the dead cells were disrupted due to cisplatin treatment. Notably, ca. 75% of genes encoding the downregulated proteins in both DL and L groups were identified as cisplatin-damaged genes (CDGs) (Fig. 4E), suggesting that cisplatin lesions are the main factor modulating protein expression.

As mentioned above, the majority of protein kinase genes and protein phosphatase genes were found to be damaged by cisplatin (Extended Data Table 3 and 4). Here, our proteomics data show that in the DL group 14 protein kinases and 4 protein phosphatases are downregulated, while in the L group 5 protein kinases and 7 protein phosphatases are downregulated due to cisplatin treatment (Fig. 4F – I). It is noteworthy that (i) more protein kinases are downregulated with high FC in the DL group than in the L group, whereas more protein phosphatases are downregulated in the L group than in the DL group, and (ii) in both DL and L group protein kinases CDK2 and BRD2 and protein phosphatases PDXP and LRRC59 are upregulated. The downregulation of the protein kinases and protein phosphatases is largely due to the cisplatin damage on the corresponding encoding genes, while the upregulation of the proteins in the cisplatin-treated cells can be attributed to the complicated cellular signaling network modulated by cisplatin damage on genes. However, the detailed mechanism underlying the upregulation of proteins including protein kinases/phosphatases requires further study.

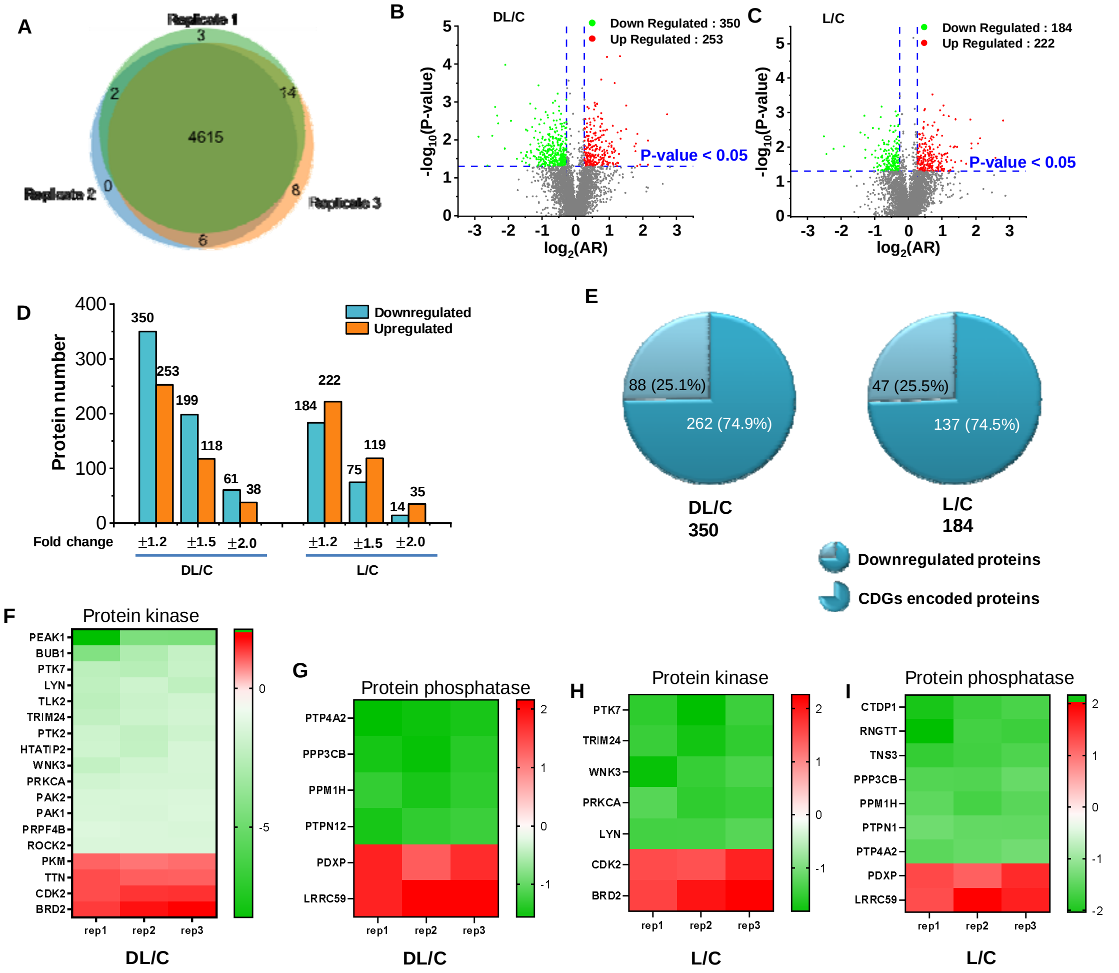

Fig. 4. Global proteomics analysis of NCCIT testicular cancer cells upon cisplatin treatment. (A) Venn diagram of total proteins identified in three biological replicates. (B, C) Volcano plots showing relative changes in protein abundance, in (B) DL group, and (C) L group, compared to the untreated cells (C group). Differentially expressed proteins (DEPs) are shown with $p < 0.05$ and fold change (FC) > 1.2 (red) or < −1.2 (green). (D) The number of DEPs detected in the DL and L groups, compared to the C group. (E) Proportion of cisplatin-damaged genes (CDGs) to the total cisplatin downregulated genes corresponding to proteins with FC values < −1.2, in the DL group and L group. (F – I) Heat maps of cisplatin-regulated protein kinases/phosphatases detected in DL and L groups.

To investigate whether cisplatin regulation to protein kinases/phosphatases disrupts and reprograms protein phosphorylation in NCCIT cells, we also carried out global phosphoproteomics on NCCIT cells subjected to cisplatin treatment (Extended Data Table 10). Compared with non-phosphorylated proteins, the abundance of phosphorylated proteins in cells is extremely low, a recognized challenge for characterization of phosphorylated proteins. [34-36] We compared the enrichment efficiency of a commercial available TiO_2 functionalized beads[37,38] and in-house synthesized ZrAs-Fe_3O_4@SiO_2 nanoparticles[39,40] for tryptic phosphopeptides derived from proteins extracted from untreated NCCIT cells. As shown in Fig. 5A, 4210 phosphopeptides were detected from the TiO_2-bead enriched sample, while 6664 phosphopeptides were detected from the ZrAs-Fe_3O_4@SiO_2 nanoparticles by nanoLC-MS. Sequential use of ZrAs-Fe_3O_4@SiO_2 nanoparticles and TiO_2 functionalized beads to enrich and isolate phosphopeptides led to detection of 11575 phosphopeptides, amongst which only 26.1% of phosphopeptides were detected in both enrichment samples (Fig. 5B). The results reveal that the sequential (or tandem) use of ZrAs-Fe_3O_4 and TiO_2 functionalized microbeads achieves a higher enrichment efficiency for phosphopeptides, compared with the use of either alone. Therefore, we used this method in the phosphoproteomic analysis for NCCIT cells treated with cisplatin. The details for sample preparation and nanoLC-MS analysis are given in experimental section and Extended Data Fig. 7A.

After preliminary database searching, we identified 11240 phosphopeptides from the three biological replicates with and without cisplatin treatment. These phosphopeptides were unambiguously assigned to 2514 phosphorylated proteins out of the 3448 proteins detected in total. The distribution of phosphorylated amino acid residues, and the proportion of different numbers of phosphorylation sites in all phosphopeptides are shown in Extended Data Fig. 7B and 7C. The numbers of the up/down regulated phosphopeptides after cisplatin treatment are shown in Fig. 5C. This high coverage of phosphopeptide data leads to identification of 601 differentially phosphorylated proteins (DPPs) with a $|FC|$ of ≥ 1.2 in DL group, and 628 DPPs in L group, compared with the control (C) group (Fig. 5D). These results indicate that cisplatin does indeed reprogram protein phosphorylation in NCCIT cells.

Next, we performed bioinformatics analysis using the String database. We assigned all DPPs as substrates to the differentially expressed protein kinases and protein phosphatases described earlier. As shown in Fig. 5E, either in DL group or L group, the majority of the DPPs are known substrates of the differentially expressed protein kinases/phosphatases. These indicate that it was the differential expression of the protein kinases/phosphatases induced by cisplatin that reprogramed protein phosphorylation in NCCIT testicular cancer cells.

Out of a total of 601 DPPs in the DL group, 41 were not assignable as substrates to any differentially expressed protein kinases (DEPKs) and/or protein phosphatases (DEPPs) as a result of cisplatin treatment. Out of 628 DPPs in the L group 30 were not assignable to any identified DEPKs and/or DEPPs. This mean that these 41 + 30 proteins are new phosphorylation substrates induced by cisplatin-induced reprogramming in NCCIT cells.

Indeed, the alignment of the amino acid sequence motifs of the phosphopeptides using the MotifeR algorithm[31] reveals the presence of new phosphorylation motifs induced by cisplatin reprogramming (Extended Data Fig. 8).

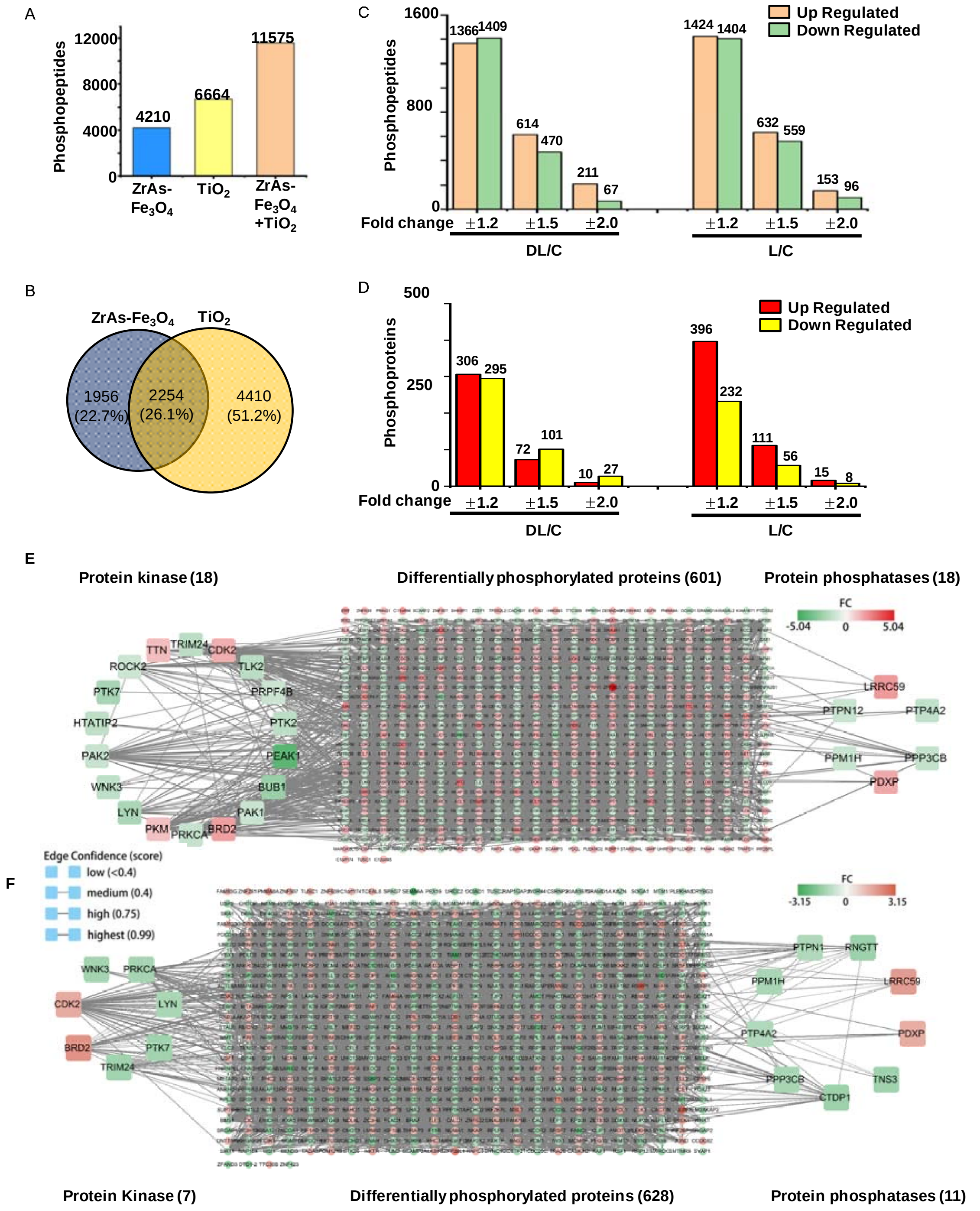

Fig. 5. Global phosphoproteomics analysis of NCCIT cells subjected to cisplatin treatment. (A) The number of phosphopeptides enriched by ZrAs-NPS, TiO_2-beads, or the sequential use of them. (B) Venn diagram of phosphopeptides enriched by either ZrAs-NPS or TiO_2-beads. (C) and (D) the numbers of phosphopeptides (C)

and phosphoproteins (D) with different fold changes for the DL group, and the L group ($p < 0.05$) identified by sequential enrichment by ZrAs-NPs and TiO_2-beads. (E – F) The phosphorylation network comprising cisplatin-induced differentially-expressed protein kinases (DEPKs) and protein phosphatases (DEPPs), and their phosphorylated/dephosphorylated proteins identified in (E) DL group, and (F) L group. In E and F, red represents upregulation and green downregulation of protein expression or protein phosphorylation.

Bioinformatics analysis of regulated proteins and phosphoproteins

To elucidate the impact of cisplatin-induced differentially-expressed proteins (DEPs) with $|FC| \geq 1.2$ and $p \leq 0.05$ on the downstream signaling pathways of NCCIT cells, bioinformatics analysis on all DEPs were carried out using Ingenuity Pathway Analysis (IPA). The top 20 signaling pathways ($p < 0.01$) are shown in Extended Data Fig. 9A and 9B, respectively. The results show that the cisplatin-induced DEPs in DL group mostly inhibit EIF2 signaling and nucleotide excision repair (NER) pathways.

The Eukaryotic initiation factor 2 (EIF2) signaling plays a crucial role in regulating translation initiation for protein synthesis[41,42]. Thus, the strong inhibition by cisplatin on EIF2 signaling could disrupt protein expression, leading to cell apoptosis and death. On the other hand, the cisplatin-inudced inhibition on NER pathway also contributes to the cell death in DL group. It is notable but not surprising that the NER pathway is predicted by IPA to be activated in the L group of NCCIT cells subjected to cisplatin treatment. This sugests that the NER activation is responsible for the survival of NCCIT cells from cisplatin treatment. In other words, the activation of NER pathway is likely one of reasons for acquired resistance against cisplatin.

IPA analysis on the differentially phosphorylated proteins (DPPs) with $|FC| \geq 1.2$ and $p \leq 0.05$ on the downstream signaling pathways of NCCIT cells indicated that in the DL group of NCCIT cells, the cisplatin-induced changes in the protein phosphorylation level mostly activate the G2/M DNA damage checkpoint regulation of cell cycle and the ATM signaling pathways (Extended Data Fig. 10). They also contribute to the inhibition of cell cycle control of chromosomal replication and the estrogen receptor signaling pathways. In the G2/M DNA Damage Checkpoint Regulation and ATM Signaling pathways, the phosphorylation level of many proteins, including the serine/threonine-protein kinase Chk2 and serine-protein kinase ATM, was upregulated or downregulated (Extended Data Fig. 11). The Chk2 protein can arrest the cell cycle by controlling important cell cycle regulators, e.g. the M-phase inducer phosphatase 3 (CDC25C), in response to DNA damage and the blocking of DNA replication. The protein encoded by the ATM gene belongs to the PI3/PI4-kinase family, which is an important cell cycle checkpoint kinase. ATM and its closely related kinase, ATR, are also major controllers of the cell cycle checkpoint signaling pathway, and required for cellular response to DNA damage and genome stability. The cisplatin-induced increase in the phosphorylation level of the proteins in the G2/M DNA damage checkpoint regulation of cell cycle and the ATM signaling pathways can activate the pathways thus leading to cell cycle arrest. Since the NER pathway in NCCIT cells is also inhibited by cisplatin, the ultimate result of cell cycle arrest by cisplatin-induced DPPs is not DNA repair, but instead, cell apoptosis due to cycle arrest at the G2/M phase. This appears to be a major contributor to the mechanism of action of cisplatin (Fig. 6).

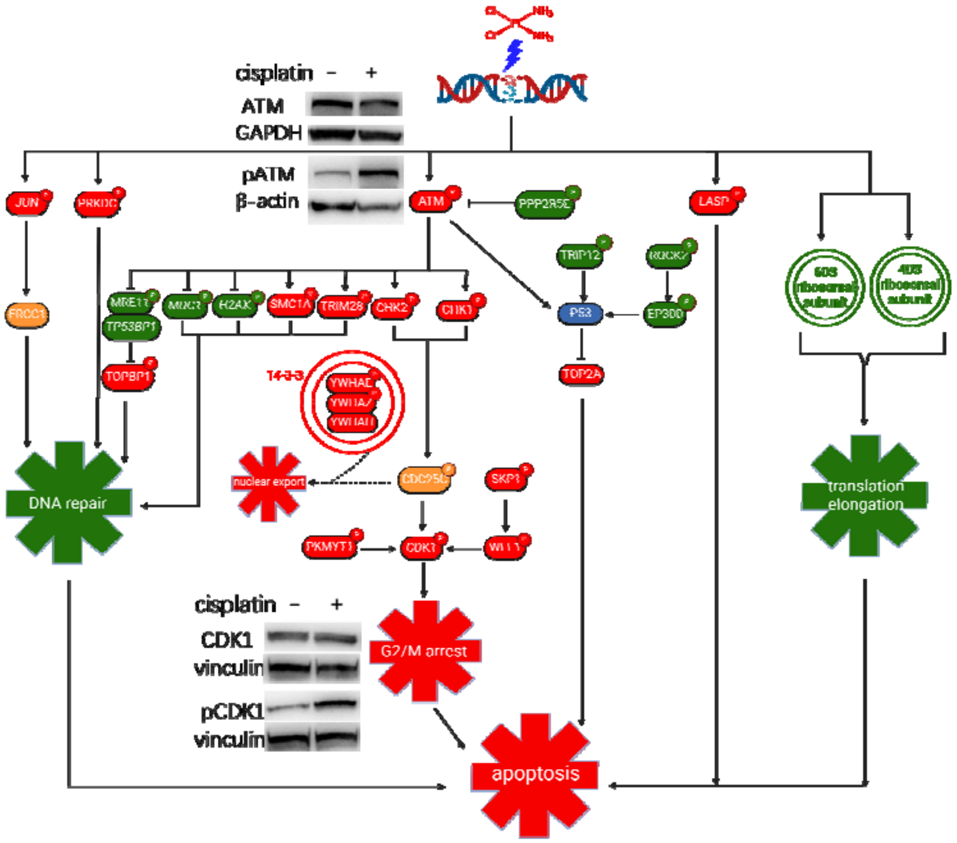
cisplatin – +
ATM
GAPDH
pATM
β-actin
JUN
PRKDC
ATM
PPP2R5E
LASP
ERCC1
MRE11
TP53BP1
MDC1
H2AX
SMC1A
TRIM28
CHK2
CHK1
TRIP12
ROCK2
P53
EP300
60S ribosomal subunit
40S ribosomal subunit
TOPBP1
14-3-3
YWHAE
YWHAZ
YWHAH
TOP2A
DNA repair
nuclear export
CDC25C
SKP1
PKMYT1
CDK1
WEE1
translation elongation
cisplatin – +
CDK1
vinculin
pCDK1
vinculin
G2/M arrest
apoptosis

Fig. 6. Cisplatin-induced reprogramming of protein phosphorylation in NCCIT cells activates G2/M DNA Damage Checkpoint Regulation and ATM signaling pathways. Red boxes indicate proteins with increased phosphorylation, and green decreased phosphorylation after cisplatin treatment. Blue box indicates IPA-predicted downregulation of the protein phosphorylation, and orange upregulation. Purple stars represent molecular functions. Arrows represent positive regulation to the phosphorylation of downstream proteins, or molecular events, and the T-type lines represent negative regulation to the phosphorylation of downstream proteins. Inserts are Western blot data for intact and phosphorylated proteins, CDK1 and ATM, subjected to cisplatin treatment.

DISCUSSION

Cisplatin is a genetically toxic anticancer drug, but its gene targets or precise gene loci are largely unknown, limiting the understanding of the cellular impact of DNA damage induced by the drug. Two methods, termed "cisplatin-seq"[12] and "damage-seq"[15], have been developed to map the cisplatin damaged sites in genomic DNA. Both these methods achieved single-base (nucleotide) resolution mapping of cisplatination sites on specific genomic regions such as promotors, transcription termination sequences and intergenic regions. However, the precise genes damaged by cisplatin and their biological consequences were not determined. Transcriptomics has been used in combination with the "damage-seq" method to map specific cisplatin damaged gene loci across mouse organs[16]. Due to the tissue-specificity of transcriptomics, only the genes transcribed in specific tissues could be mapped for cisplatin lesions.

There are a number of chemical genetic strategies, in particular chemical proteomic approaches, that can uncover gene targets for cisplatin. These have produced significant advances in elucidating mechanisms of drug action. However, similar to the transcriptomics approach, these studies do not fully reveal cisplatin-damaged gene loci throughout the whole genome because they focus on discovering gene targets of cisplatin *via* measuring the difference between gene expression in specific types of cancer cells or tumour tissues upon cisplatin exposure. For example, Jimenez and co-workers found that UGGT1, COL6A1 and MAP4 can be used as potential biomarkers for determining cisplatin response, by proteomic analysis of non-small cell lung cancer cells[43]. Also Wilmes et al. have used integrated omics techniques to study differentially-expressed proteins and related pathways for cisplatin nephrotoxicity[44].

In this work we have used the principle of forward chemical genetics[17,18], combining affinity separation and high throughput sequencing to map cisplatin-damaged gene loci and their influence on cellular signalling pathways. Importantly, by integration with bioinformatics analysis, our data provide three important and innovative findings. Firstly, amongst the 11952 identified cisplatin-damaged protein-encoding genes, 418 genes are in the 628 prior gene targets for cancer treatment reported recently[26] (Extended Data Table 5). The second key discovery is that mitochondrial genes are not the major targets for cisplatin. There are two possible reasons, one that mitochondrial genes are usually short, and also have a low density of guanine. This result is in consistent with the widely accepted knowledge that nuclear DNA is the major target of cisplatin[3,5]. Thirdly, and importantly cisplatin attacks 400 protein kinase genes with fold-enrichment >1.5 (Extended Data Table 3), covering 77% of consensual protein kinase genes in the human genome[25]. Moreover, cisplatin also induced damage on 131 protein phosphatase genes, accounting for 96% of the 136 reported protein phosphatases in human genome (Extended Data Table 4). From these findings it is clear that cisplatin-damaged genes significantly disturb the signalling of protein phosphorylation in cells, with the potential to reprogram protein phosphorylation proteome-wide.

Our proteomics and phosphoproteomics studies for human testicular cancer (NCCIT) cells revealed that cisplatin treatment down/up-regulated many important protein kinases and protein phosphatases, which indeed reprograms protein phosphorylation in NCCIT cells (Extended Data Table 9 and 10). Significantly, cisplatin-induced protein phosphorylation reprogramming activates the G2/M DNA damage checkpoint regulation and ATM signalling pathways, leading to cell cycle arrest and ultimately to cell apoptosis and death.

Protein kinases accompanied by phosphatases regulate protein phosphorylation, the most common and important protein post-translational modification, involved in many natural biochemical pathways and cellular processes[45]. Since abnormal protein phosphorylation is usually and closely related to cancer development, protein kinases are major drug targets for cancer therapy[45,46]. A number of small molecule protein kinase inhibitors (PKIs), e.g. gefitinib, imatinib and dasatinib[47], have been developed and approved for clinical treatment of various cancers. Given our findings that protein kinase/phosphatase genes are highly susceptible to cisplatin attack, this work demonstrates that cisplatin can control/regulate broad-spectrum protein phosphorylation by inducing lesions on protein kinase and protein phosphatase genes. Pharmacological functions of cisplatin may therefore involve genome-wide reprograming of protein phosphorylation in cancer cells. Our findings provide a basis for the design of next-generation platinum and other metal anticancer drugs, or drug delivery systems, to minimize non-specific (undesired) attack on protein kinase/phosphatase genes which are essential for healthy cells.[2,48]

REFERENCES

1 Ramalingam, S. S., Owonikoko, T. K. & Khuri, F. R. Lung cancer: new biological insights and recent therapeutic advances. *CA Cancer J. Clin.* **61**, 91-112, (2011).

2 Johnstone, T. C., Suntharalingam, K. & Lippard, S. J. The next generation of platinum drugs: targeted Pt(II) agents, nanoparticle delivery, and Pt(IV) prodrugs. *Chem. Rev.* **116**, 3436-3486, (2016).

3 Jamieson, E. R. & Lippard, S. J. Structure, recognition, and processing of cisplatin-DNA adducts. *Chem. Rev.* **99**, 2467-2498, (1999).

4 Mezencev, R. Interactions of cisplatin with non-DNA targets and their influence on anticancer activity and drug toxicity: the complex world of the platinum complex. *Curr. Cancer Drug Tar.* **14**, 794-816, (2014).

5 Jung, Y. & Lippard, S. J. Direct cellular responses to platinum-induced DNA damage. *Chem. Rev.* **107**, 1387-1407, (2007).

6 Yang, Z. *et al.* Cisplatin Preferentially Binds Mitochondrial DNA and Voltage-Dependent Anion Channel Protein in the Mitochondrial Membrane of Head and Neck Squamous Cell Carcinoma: Possible Role in Apoptosis. *Clin. Cancer Res.* **12**, 5817-5825, (2006).

7 Qi, L. *et al.* Advances in toxicological research of the anticancer drug cisplatin. *Chem. Res. Toxicol.* **32**, 1469-1486, (2019).

8 Zeng, W. *et al.* Proteomic strategy for identification of proteins responding to cisplatin-damaged DNA. *Anal. Chem.* **91**, 6035-6042, (2019).

9 Ohndorf, U. M., Rould, M. A., He, Q., Pabo, C. O. & Lippard, S. J. Basis for recognition of cisplatin-modified DNA by high-mobility-group proteins. *Nature* **399**, 708-712, (1999).

10 Jia, F. F. *et al.* In Situ Visualization of Proteins in Single Cells by Time-of-Flight-Secondary Ion Mass Spectrometry Coupled with Genetically Encoded Chemical Tags. *Anal. Chem.* **92**, 15517-15525, (2020).

11 Lin, Y. *et al.* Single cell imaging reveals cisplatin regulating interactions between transcription (co)factors and DNA. *Chem. Sci.* **12**, 5419-5429, (2021).

12 Shu, X., Xiong, X., Song, J., He, C. & Yi, C. Base-resolution analysis of cisplatin-DNA adducts at the genome scale. *Angew. Chem. Int. Ed.* **55**, 14244-14247, (2016).

13 Guo, W. *et al.* Transferrin serves as a mediator to deliver organometallic ruthenium(II) anticancer complexes into cells. *Inorg. Chem.* **52**, 5328-5338, (2013).

14 Mingard, C., Wu, J., McKeague, M. & Sturla, S. J. Next-generation DNA damage sequencing. *Chem. Soc. Rev.* **49**, 7354-7377, (2020).

15 Hu, J., Lieb, J. D., Sancar, A. & Adar, S. Cisplatin DNA damage and repair maps of the human genome at single-nucleotide resolution. *Proc. Natl. Acad. Sci. USA* **113**, 11507-11512, (2016).

16 Yimit, A., Adebali, O., Sancar, A. & Jiang, Y. Differential damage and repair of DNA-adducts induced by anti-cancer drug cisplatin across mouse organs. *Nat. Commun.* **10**, 309, (2019).

17 Walsh, D. P. & Chang, Y.-T. Chemical genetics. *Chem. Rev.* **106**, 2476-2530, (2006).

18 Johnson, E. O. *et al.* Large-scale chemical-genetics yields new M. tuberculosis inhibitor classes. *Nature* **571**, 72-78, (2019).

19 Liu, X., Bing, T. & Shangguan, D. Microbead-Based Platform for Multiplex Detection of DNA and Protein. *ACS Appl. Mater. Inter.* **9**, 9462-9469, (2017).

20 Wang, J. *et al.* Direct Quantification of MicroRNA at Low Picomolar Level in Sera of Glioma Patients Using a Competitive Hybridization Followed by Amplified Voltammetric Detection. *Anal. Chem.* **84**, 6400-6406, (2012).

21 Meng, X., Sun, P., Xu, H. & Wang, Z. Folic acid-functionalized magnetic nanoprobes via a PAMAM dendrimer/SA-biotin mediated cascade-amplifying system for the efficient enrichment of circulating tumor cells. *Biomater. Sci.* **8**, 6395-6403, (2020).

22 Onoa, G. B., Cervantes, G., Moreno, V. & Prieto, M. J. Study of the interaction of DNA with cisplatin and other Pd(II) and Pt(II) complexes by atomic force microscopy. *Nucleic Acids Res.* **26**, 1473-1480, (1998).

23 Marini, V. *et al.* Conformation, protein recognition and repair of DNA interstrand and intrastrand cross-links of antitumor trans-[PtCl2(NH3)(thiazole)]. *Nucleic Acids Res.* **33**, 5819-5828, (2005).

24 Reedijk, J. Why does cisplatin reach guanine-N7 with competing S-donor ligands available in the cell? *Chem. Rev.* **99**, 2499-2510, (1999).

25 Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The protein kinase complement of the human genome. *Science* **298**, 1912-1934, (2002).

26 Behan, F. M. *et al.* Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens. *Nature* **568**, 511-516, (2019).

27 Lou, W., Ding, B. & Fu, P. Pseudogene-Derived lncRNAs and Their miRNA Sponging Mechanism in Human Cancer. *Front. Cell Dev. Bio.* **8**, (2020).

28 Cheetham, S. W., Faulkner, G. J. & Dinger, M. E. Overcoming challenges and dogmas to understand the functions of pseudogenes. *Nat. Rev. Genet.* **21**, 191-201, (2020).

29 Yan, Q. *et al.* SPAG9 is involved in hepatocarcinoma cell migration and invasion via modulation of ELK1 expression. *OncoTargets Ther.* **9**, 1067-1075, (2016).

30 Kelkar, N., Standen, C. L. & Davis, R. J. Role of the JIP4 scaffold protein in the regulation of mitogen-activated protein kinase signaling pathways. *Mol. Cell. Bio.* **25**, 2733-2743, (2005).

31 Jiang, Y. *et al.* MERTK mediated novel site Akt phosphorylation alleviates SAV1 suppression. *Nat. Commun.* **10**, 1515, (2019).

32 Graham, D. K., DeRyckere, D., Davies, K. D. & Earp, H. S. The TAM family: phosphatidylserine-sensing receptor tyrosine kinases gone awry in cancer. *Nat. Rev. Cancer* **14**, 769-785, (2014).

33 Khan, P. *et al.* Aspirin enhances cisplatin sensitivity of resistant non-small cell lung carcinoma stem-like cells by targeting mTOR-Akt axis to repress migration. *Sci. Rep.* **9**, 16913, (2019).

34 Ochoa, D. *et al.* The functional landscape of the human phosphoproteome. *Nat. Biotechnol.* **38**, 365-373, (2020).

35 Tsai, C.-F. *et al.* Large-scale determination of absolute phosphorylation stoichiometries in human cells by motif-targeting quantitative proteomics. *Nat. Commun.* **6**, 6622, (2015).

36 Wu, R. *et al.* A large-scale method to measure absolute protein phosphorylation stoichiometries. *Nat. Methods* **8**, 677-683, (2011).

37 Ji, L. *et al.* Quantitative Mass Spectrometry Combined with Separation and Enrichment of Phosphopeptides by Titania Coated Magnetic Mesoporous Silica Microspheres for Screening of Protein Kinase Inhibitors. *Anal. Chem.* **84**, 2284-2291, (2012).

38 Zhai, G. *et al.* Evaluation of serum phosphopeptides as potential cancer biomarkers by mass spectrometric absolute quantification. *Talanta* **125**, 411-417, (2014).

39 Zhai, G. *et al.* Serum phosphopeptide profiling for colorectal cancer diagnosis using liquid chromatography–mass spectrometry. *Rapid Commun. Mass Spectrom.* **36**, e9316, (2022).

40 Zhai, G. *et al.* Evaluation of serum phosphopeptides as potential biomarkers of gastric cancer. *RSC Advances* **7**, 21630-21637, (2017).

41 Lee, A. S. Y., Kranzusch, P. J. & Cate, J. H. D. eIF3 targets cell-proliferation messenger RNAs for translational activation or repression. *Nature* **522**, 111-114, (2015).

42 Masutani, M., Sonenberg, N., Yokoyama, S. & Imataka, H. Reconstitution reveals the functional core of mammalian eIF3. *EMBO J.* **26**, 3373-3383, (2007).

43 Bottger, F. *et al.* Proteome analysis of non-small cell lung cancer cell line secretomes and patient sputum reveals biofluid biomarker candidates for cisplatin response prediction. *J. Proteome.* **196**, 106-119, (2019).

44 Wilmes, A. *et al.* Mechanism of cisplatin proximal tubule toxicity revealed by integrating transcriptomics, proteomics, metabolomics and biokinetics. *Toxicology in Vitro* **30**, 117-127, (2015).

45 Cohen, P. Protein kinases - the major drug targets of the twenty-first century? *Nat. Rev. Drug. Discov.* **1**, 309-315, (2002).

46 Katsogiannou, M., Andrieu, C. & Rocchi, P. Heat shock protein 27 phosphorylation state is associated with cancer progression. *Front. Genet.* **5**, 346, (2014).

47 Baselga, J. Targeting tyrosine kinases in cancer: The second wave. *Science* **312**, 1175-1178, (2006).

48 Wang, X. & Guo, Z. Targeting and delivery of platinum-based anticancer drugs. *Chem. Soc. Rev.* **42**, 202-224, (2013).

49 Park, P. J. ChIP-seq: advantages and challenges of a maturing technology. *Nat. Rev. Genet.* **10**, 669-680, (2009).

50 Kumar, D., Mariappan, G., Husain, A., Monga, J. & Kumar, S. Design, synthesis and cytotoxic evaluation of novel imidazolone fused quinazolinone derivatives. *Arab. J. Chem.* **10**, 344-350, (2017).

51 Mertins, P. *et al.* Reproducible workflow for multiplexed deep-scale proteome and phosphoproteome analysis of tumor tissues by liquid chromatography–mass spectrometry. *Nat. Protoc.* **13**, 1632-1661, (2018).

52 Kraemer, A., Green, J., Pollard, J., Jr. & Tugendreich, S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* **30**, 523-530, (2014).

Materials and Methods

Escherichia coil

E. coli strain BL21 with HMGB1a-encoding plasmid (DE3) (Tian Gen Biotech, China) was cultivated in Luria-Bertani (LB) medium with 50 μg/mL kanamycin. A single colony picked from a fresh plate was used to inoculate 2YT medium (containing Bacto^{TM} Peptone (16 mg/mL), Bacto^{TM} Yeast Extract (10 mg/mL, Becton, Dickinson and company, USA), sodium chloride (5 mg/mL, Beijing Chemical Works, China) and kanamycin (50 μg/mL, Solarbio, China)) overnight at 37 °C and 180 rpm. Then the cells were transferred to a large volume of fresh 2YT medium at 37 °C and 180 rpm. When the OD_{600} value of the cells reached 0.5 – 1.0, IPTG (Sigma-Aldrich, USA) was added at a final concentration of 1 mM to induce gene expression. After incubation at 37 °C and 180 rpm for further 12 h, the cells were harvested and washed once with PBS (Solarbio, China), and frozen at -80°C for at least one hour for better sonication.

Cell lines

The A549 human non-small cell lung cancer cell line, HeLa cervical cancer cell line, HepG2 hepatoma cell line (National Infrastructure of Cell Line Resource, Beijing, China) and NCCIT testicular cancer cell line (Fukesai Biological Technology Co., Ltd, Nanjing, China) were individually grown in DMEM medium (Gibco) supplemented with 10% FBS (Gibco) and 1% penicillin and streptomycin (GE Healthcare) at 37 °C. When the cell density reached confluency, the culture medium was replaced by fresh DMEM in the absence or presence of cisplatin at various concentrations (50 μM for A549 and HepG2, 12 μM for NCCIT, 25 μM for HeLa) for 24 h. Then the cells were harvested and washed twice by PBS. Whole cell protein or genomic DNA were extracted from different cells with different extraction kits.

Expression and purification of HMGB1a

The cDNA of human HMGB1 domain A (HMGB1a, Extended Data Fig. 12A) was synthesized by Proteintech (Rosemont, USA) and cloned into pET-28a plasmid (Extended Data Fig. 12B). The resulting plasmid encoding a his_6-tag was transfected into BL21 (DE3) *E. coli* cells (Tian Gen Biotech, China). The cells were grown in a small volume of 2YT medium containing Bacto^{TM} Peptone (16 mg/mL), Bacto^{TM} Yeast Extract (10 mg/mL, Becton, Dickinson and company, USA), sodium chloride (5 mg/mL, Beijing Chemical Works, China) and kanamycin (50 μg/mL, Solarbio, China) overnight at 37 °C and 180 rpm. Then the cells were transferred to a large volume of fresh 2YT medium at 37 °C and 180 rpm. When the OD_{600} value of the cells reached 0.5 – 1.0, IPTG (Sigma-Aldrich, USA) was added at a final concentration of 1 mM to induce gene expression. After incubation at 37 °C and 180 rpm for further 12 h, the cells were harvested and washed once with PBS (Solarbio, China), and frozen at -80°C for at least one hour for better sonication. The deep frozen cells were resuspended in 10 mL cold lysis buffer (PBS and 5 mM imidazole, pH 7.4), and lysed by sonication for 20 min at 0 °C, followed by centrifugation for 30 min at 4 °C and 12000 rpm. The supernatant was purified by His Trap^{TM} FF crude column (GE Healthcare) which was eluted first by 50 mM imidazole (Innochem), then by 300 mM imidazole in phosphate buffer. The eluted HMGB1a containing a his_6-tag was characterized by SDS-PAGE (Extended Data Fig. 12C) and mass spectrometry (Extended Data Fig. 12D).

Assembly and characterization of HMGB1a-functionalized affinity microprobes

To construct affinity microprobes for cisplatin-crosslinked DNA, 2.1 mg purified HMGB1a was incubated with 7.5 mg nickel-modified magnetic beads (MAg25K/IDA Ni, Enriching Biotechnology, Shanghai, China) suspended in 1 mL PBS, and rotated for 1 h at 4 °C. The beads were then separated by magnetic force and washed twice with PBS for 10 min. As shown in Extended Data Fig. 12E, the absorbance of the supernatant over the wavelength range 250 – 300 nm significantly decreased due to loading of HMGB1a onto the magnetic beads. This indicated successful functionalization of the magnetic beads by HMGB1a.

Cell culture and genomic DNA extraction

The A549 human non-small cell lung cancer cells obtained from the Centre for Cell Resource of Peking Union Medical College Hospital were cultured in DMEM medium (Gibco) supplemented with 10% FBS (Gibco) and 1% penicillin and streptomycin (GE Healthcare) at 37 °C. When the cell density reached confluency, the culture medium was replaced by fresh DMEM containing 50 μM cisplatin (Beijing Ouhe Technology, China), and the cells were incubated at 37 °C for 24 h. Then the cells (~10^7) were harvested, and genomic DNA was extracted by whole genome extraction kit (Tian Gen Biotech). The DNA extracts were crushed into fragments between 250 to 750 bp with an ultrasonic breaker (Scientz-IID ultrasonic homogenizer, SCIENTZ, Ningbo, China). At the same time, a similar amount of cells without exposure to cisplatin were harvested for genomic DNA extract and fragmentation. The DNA fragments from the cells treated with cisplatin were considered as a positive sample and those from the cells without cisplatin treatment as a control sample for capture and sequencing of cisplatin damaged genes. An inductively coupled plasma-mass spectrometer (ICP-MS, Agilent 7700) was used to determine the Pt concentrations of the DNA fragments obtained from positive and control samples.

Capture of cisplatin-bound DNA fragments

The same amounts (74 μg) of DNA fragments gained from cells treated and untreated with cisplatin were individually re-dissolved in 10 mM Tris-HCl (pH 7.4, Beijing Biodee Biotechnology), and added into the HMGB1a-functionlized magnetic beads suspended in PBS, followed by rotation for 1 h at room temperature. The beads were then separated by magnetic force and washed twice with 1 mL 10 mM Tris-HCl. The captured HMGB1a-DNA complexes were eluted with 1 mL of buffer containing 20 mM sodium phosphate, 500 mM NaCl and 400 mM imidazole (pH 7.4) *via* rotation at room temperature for 40 min. As shown in Extended Data Fig. 12F, G, the absorbance of the supernatant over the wavelength range 250 – 300 nm significantly decreased subject to incubation with HMGB1a-functionlized magnetic beads, providing evidence for the capture of DNA fragments by the HMGB1a-based affinity microprobes.

Each eluted HMGB1a-DNA complex sample was centrifuged and lyophilized, and then re-dissolved in 500 μL deionized water, followed by addition of 500 μL mixture of phenol (Solarbio) – chloroform (Beijing Chemical Works) – isopropanol (Concord Technology, Tianjin) (25: 24: 1). The resulting mixture was shaken firstly for 10 s, and kept at room temperature for 2 min, then centrifuged at 12000 rpm and 4 °C for 7 min. The supernatant was mixed with 500 μL of a mixture of chloroform-isopropanol (24: 1), shaken again for 10 s, kept at room temperature for 2 min, and then centrifuged at 12000 rpm and 4 °C for 7 min. The isolated supernatant was mixed with 1500 μL pre-chilled absolute ethanol (Concord Technology, Tianjin) and 50 μL sodium acetate solution (3 M, Beijing Chemical Works), kept at −20 °C overnight, and then centrifuged at 12000 rpm and 4 °C for 7 min. The supernatant was carefully removed, and the residues were left at room temperature for drying.

Finally, the purified DNA fragments were re-dissolved in an appropriate amount of TE solution (Solarbio) for sequencing.

NGS Sequencing and gene mapping

The next-generation-sequencing (NGS) of the captured and purified DNA fragments described above was performed by Shanghai Sangon Biotech Ltd. The DNA fragments were firstly characterized by 1% agarose gel electrophoresis for quality check, and then fragmented to about 500 bp by S220 Focused-ultrasonicator and recovered for DNA library construction. Library construction was performed following the instruction of NEB Next® Ultra™ DNA Library Prep Kit for Illumina® (https://international.neb.com/products/e7370-nebnext-ultra-dna-library-prep-kit-for-illumina#Protocols,%20Manuals%20&%20Usage). Since cisplatin binding to DNA fragments inhibits DNA replication and synthesis[3,5,23], the PCR expansion time was prolonged to 30 min. The constructed DNA libraries were characterized by a 2% agarose gel, and then each positive or control sample was sequenced by Illumina®, providing at least 20 G of data.

The raw Illumina sequences were filtered for quality control by using Fastp (0.22.0) (fastp -q 15 -u 40 -n 5 -e 0 -l 35), and the clean reads were aligned to the reference human genome (hg19; National Centre for Biotechnology Information [NCBI]) by using the Burrows-Wheeler Aligner(BWA) software program (0.7.17-r1188) with default parameter [bwa mem -M <genome> <R1> <R2>] to acquire SAM files. Then, Samtools (1.6) was used to sort out the SAM files to BAM files in which only uniquely mapped nonduplicate reads were retained. High resolution ChIP-Seq peaks predictions were conducted by MACS2(2.1.2) program with a p-value cutoff of 0.001 and Mfold values adjusted to build the model, [macs2 callpeak -c <treat BAM > -t <control BAM> -q 0.05 -f BAMPE -g hs]. Finally, the R-Chipseeker package was employed to annotate and visualize all peaks identified.

To calculate the fold-enrichment of a cisplatin damaged gene (CDG), we first calculated the fold-enrichment of all peaks mapped to the gene by using equation 1.

$$\mathrm{FE_p} = \frac{\mathrm{R_{PS}} / T_{PS}}{\mathrm{R_{CS}} / T_{CS}} \tag{1}$$

Where R_{PS} and R_{CS} are read (tag) counts for a peak detected in a pair of positive sample (PS) and control sample (CS), respectively, and T_{PS} and T_{CS} are total read (tag) counts of all peaks detected in the pair of positive sample (PS) and control sample (CS), respectively. A peak was considered unreliable and then abandoned if it matched multiple genes. A peak with an fold-enrichment (FE_p) <1.5 was also taken as unreliable and discarded[49]. Thus, the fold-enrichment of a CDG is the sum of fold-enrichment of all peaks mapped to the gene, equation 2.

$$\mathrm{FE_G} = \sum_{i}^{i} (FE_p)_i \tag{2}$$

Where i represents the number of peaks mapped for a gene in a positive sample/control sample pair.

Taking into account that cisplatin kills cells by attack on DNA, it might be expected that more cisplatin-damaged genes would be detected in the dead cells of positive samples. However, our sequencing data showed that there was no significant difference between the total numbers and fold-enrichments of cisplatin-damaged genes mapped from the DNA fragments collected from living cells alone and from living cells and dead cells together after 24 h of incubation with 50 μM cisplatin (data not shown). The other software and databases used for gene screening and localization are listed in Extended Data Table 11.

CCK-8 assay

The survival rates of *SPAG9*-selicenced and cisplatin-treated NCCIT cells were measured using the CCK-8 (Cell Counting Kit-8) assay (MedChemExpress, USA). Briefly, the NCCIT cells ($1×10^3$ cells/per wells) were plated at a density of 2000 cells per well in 100 μL of media into 96-well plates. After 24 h for adherence, the cells were exposed to 12 μM cisplatin for 24 h or transfected with the siRNA against *SPAG9* and grown for 48 h. Then, 10 μL CCK-8 was added into the medium and the cells were grown for further 4 h. Finally, the optical density (OD) for each well was measured using a microplate reader (SpectraMax M5 Molecular Devices Corporation) at the wavelength of 450 nm. The survival rate (SR, %) of cells was calculated based on the equation (1):

$$SR(\%) = \frac{OD_{sample} - OD_{blank}}{OD_{control} - OD_{blank}} \times 100\% \qquad (3)$$

All reported values of SR were averages of three independent experiments and expressed as mean ± SD (standard deviation).

Flow cytometric analysis

The apoptosis rate of NCCIT cells subject to cisplatin treatment and *SPAG9* silencing was determined by Annexin V / propidium iodide (PI) double staining assay following the Apoptosis Detection Kit I protocols (BD Pharmingen, USA). The NCCIT cells ($1×10^3$ cells/well) were plated into a six-well plate. Twenty-four hours later, NCCIT cells were transfected with siRNA against *SPAG9* for 48 h. Another group of cells were treated with 12 μM cisplatin for 24 h. Then, the supernatant was removed, and cells were transferred to tubes and centrifuged at 1000 rpm for 3 min. Thereafter, the cells were re-suspended in 1× Annexin V binding buffer, and incubated with 5 μL Annexin-V and 5 μL PI conjugate for 15 min in the dark prior to FACS analysis. The FACS assays were performed on a Calibur flow cytometer (BD , Franklin Lakes, New Jersey, US), of which the FL1 channel was used to record the intensity of annexin V-FITC staining and FL2 channel to record the intensity of PI staining. The data were quantified by Sell Quest software (BD, Franklin Lakes, New Jersey, US).

Quantification of protein and DNA concentration

The concentrations of purified protein HMGB1a and extracted genomic DNA were determined by NanoDrop OneC (Thermo Fisher Scientific Inc.). The concentrations of whole cell proteins from different cells were determined using a BCA kit (Beyotime Biotech Inc.).

Western Blot assays

The A549 human non-small cell lung cancer cell line, HeLa cervical cancer cell line, HepG2 hepatoma cell line (National Infrastructure of Cell Line Resource, Beijing, China) and NCCIT testicular cancer cell line (Fukesai Biological Technology Co., Ltd, Nanjing, China) were individually grown in the presence and absence of cisplatin at various concentrations (50 μM for A549 and HepG2, 12 μM for NCCIT, 25 μM for HeLa) for 24 h. The cells were then harvested for extracting whole cell proteins using a total protein extraction kit (BestBio, China). The concentration of each protein extract was determined by standard BCA assay (Beyotime, China). To determine the change in expression level of a specific protein due to gene damage induced by cisplatin, the same amount of whole cell proteins extracted from a positive sample and a control sample were individually boiled with the gel-loading buffer (4×LDS, Genscript, Nanjing, China), and loaded onto a 4 – 12% gradient SDS-PAGE gel (Genscript) for electrophoresis. The separated proteins were transferred to a PVDF membrane (Millipore, 0.2 μm). The membrane was blocked in 5% nonfat dry milk (Biofroxx) dissolved in 0.1% TBST, of which 1 L contained 2.4 g Tris base (Innochem), 8.8 g NaCl and 0.1% (v/v) Tween 20 (Beijing Topbio Science&Technology), at 4 °C overnight, and then incubated with respective primary antibodies (Abcam) in appropriate dilutions at room temperature for 2 h. Thereafter, the membranes were washed three times with 5% nonfat dry milk in 0.1% TBST for 30 min each time and incubated with horseradish peroxidase-conjugated secondary antibody (Abcam) at room temperature for 1 h. After washing twice with 5% nonfat dry milk in 0.1% TBST and once with 0.1% TBST, the protein bands were visualized by enhanced chemiluminescence (Beyotime, China), and the optical densities were determined using an image analyser (Image J).

Since these chosen genes are expressed at different levels in different types of cells, we selected four human cancer cell lines, A549 (non-small cell lung cancer cell line), NCCIT (testicular cancer cell line), HepG2 (hepatocellular cancer cell line) and HeLa (cervical cancer cell line) to perform Western Blot assays. Among the four cell lines, NCCIT is the most sensitive to cisplatin with an IC_{50} value of 5.8 μM (24 h treatment). The IC_{50} values for HeLa, A549 and HepG2 cells were ca. 8.2 μM (24 h), 10.1 μM (48 h), 22.4 μM (24 h)[50], respectively. Thus, we applied different doses of cisplatin to treat the cell lines, i.e. 50 μM for A549 and HepG2, 25 μM for HeLa and 12 μM for NCCIT.

For siRNA silencing of *SPAG9* experiments, the NCCIT testicular cancer cell line were transfected with siRNA against *SPAG9* (sequence 5'-3': rGrGrArGrCrArGrArUrUrUrArCrUrArGrGrArATT, Suzhou Biosyn Biotechnology Co., Ltd.), or not, for 48 h. The cells were then harvested for extracting whole cell proteins using a total protein extraction kit (BestBio, China). The concentration of each protein extract was determined by standard BCA assay (Beyotime, China). To determine the change in expression level of JIP-4 due to gene silencing by siRNA, the same amount of whole cell proteins extracted from a positive sample and a control sample were individually analysed by using the same method for Western Blot assays described above.

Calculation of fold-change of protein expression

We carried out three technical repetitions of Western Blot assays. The optical densities of protein bands were determined using an image analyser (Image J). The fold-change (FC) of a protein was calculated using equation 4.

$$\mathrm{FC} = \frac{\mathrm{OD_{PC}/OD_{GC}}}{\mathrm{OD_{PN}/OD_{GN}}} \tag{4}$$

Where OD_{PC} and OD_{GC} are optical densities of a protein band and the GAPDH (the reference) band in a sample treated with cisplatin, and OD_{PN} and OD_{GN} are optical densities of the protein band and the GAPDH band in a sample without cisplatin treatment, respectively.

Proteomics/Phosphoproteomics

Cell culture and protein extraction

The NCCIT testicular cancer cell line obtained from Fukesai Biological Technology Co., Ltd were cultured in RPMI-1640 medium (Gibco) supplemented with 10% FBS (Gibco) and 1% penicillin and streptomycin (GE Healthcare) at 37°C. When the cell density reached confluency, the culture medium was replaced by fresh RPMI-1640 containing 12 μM cisplatin (Beijing Ouhe Technology, China), and the cells were incubated at 37 °C for 24 h. Cells without cisplatin treatment were set as C group. A group of cells were treated with cisplatin for 24 h and only live cells were collected as L/C group. Another group of cells were treated with cisplatin for 24 h and both live cells and dead cells were collected as DL/C group. The cells were harvested for extracting whole cell proteins using a total protein extraction kit (BestBio, China) supplemented with protease inhibitor, phosphatase inhibitor and 8 M urea. The concentration of each protein extract was determined by standard BCA assay (Beyotime, China).

Sample preparation

Reduce the disulfide bonds in the denatured proteins by adding DTT (J&K Scientific) to 300 μg (for phosphoproteomics, the amount of initial protein is 1 mg) of sample (group C, L and DL, three biological replicates) at a final concentration of 5 mM DTT, followed by incubation at 37 °C for 1 h in a Thermomixer. Then alkylate the reduced cysteine residues by adding IAA (J&K Scientific) to a final concentration of 10 mM IAA, followed by incubation for 45 min in the dark at 25 °C. Dilute the sample 1:3 (vol/vol) with 50 mM Tris-HCl (pH 8.0, Beijing BioDee Biological Technology Co., Ltd) to decrease the urea concentration to <2 M. Add trypsin (Sigma-Aldrich) in an enzyme/substrate ratio of 1:49 (wt/wt) for overnight digestion at 25 °C in the incubator shaker. Then quench the digestion reaction by acidifying the digestion mixture to a final concentration of 1% (vol/vol) formic acid (FA) using 100% (vol/vol) FA and dilute the samples with 0.1% (vol/vol) FA to a 1.5-mLvolume, followed by desalting by C18 cartridge (Waters Sep-Pak). The C18 cartridge was firstly activated by 1 mL of 100% acetonitrile (ACN, Fisher Chemical) and 1 mL of 0.1% FA/50% ACN sequentially, and then equilibrated by 1 mL of 0.1% trifluoroacetic acid (TFA, Alfa Aesar) for three times. The collected peptides were loaded into the equilibrated column, and washed by 1 mL of 0.1% TFA for three times and 1 mL of 1% FA, and then eluted by 1 mL of 0.1% FA/50% ACN and 1 mL of 0.1% FA/80% ACN sequentially. Finally, the desalted peptides were merged and dried in vacuum centrifuge for further use.

TMT labeling

Reconstitute the sample in 100 μL of 50 mM HEPES (pH 8.5) by mixing well. Remove a 5-μL aliquot for the peptide-level BCA estimation. On the basis of the BCA results at the peptide level, dilute the samples with 50 mM HEPES (pH 8.5) for a concentration of 1 μg/μL. Vortex the well to mix. Reconstitute each 0.8 mg TMT reagent (Thermo Fisher, for phosphoproteomics, the amount of TMT reagent is 5 mg) vial in 41 μL of anhydrous acetronitrile. Allow it to sit for 5 min. Vortex the reagent tubes briefly and spin down on a benchtop centrifuge at 1000 g at RT for 30 s. Add 41 µL of each reagent to the corresponding aliquot of peptides. Mix well and incubate at 25 °C for 1 h with shaking on a tabletop shaker set at 300 rpm. After 1 h, quench all individual TMT-labeling reactions by adding 8 µL of 5% (vol/vol) hydroxylamine (J&K Scientific), followed by incubation at RT for 15 min. Combine the volumes of each of the nine TMT-labeled samples in a 1.5 mL conical tube and add 0.1% FA to 1.5 mL. Then desalted this sample by using the method as described above. The desalted peptides were merged and dried in vacuum centrifuge for further use.

High-pH reversed-phase pre-separation of peptides

Prepare the mobile phases A and B required for pre-separation according to the literature[51]. The TMT-labeled and desalted peptide mixture was re-dissolved in 100 µL mobile phases A. Aliquot (95 µL) of each labeled peptide mixture was loaded into HPLC (Agilent Technologies 1260 infinity) with a C18 column (CAPCELL PAK MG II S-5, 4.6 × 150 mm, 5 µm, SHISEIDO), and eluted by basic (pH =10) mobile phases A (4.5 mM $HCOONH_4$ in 2% ACN) and B (4.5 mM $HCOONH_4$ in 90% ACN). The gradient started with 0% B until 7 min and continuously increased to 16% B at 13 min, 40% B at 73 min, 44% B at 77 min, 60% B at 82 min and kept until 96 min with a flow rate of 1 mL/min. The fractions were collected chronologically into twelve tubes from 3 min to 90 min, individually lyophilized and re-dissolved in 10 µL of 0.1% formic acid for NanoLC-MS/MS analysis.

Phosphopeptide enrichment

The twelve peptide solutions obtained by pre-separation were dried in vacuum centrifuge separately, and then continuously enriched by using ZrAs-Fe_3O_4@SiO_2 magnetic nanoparticles (ZrAs-NPs)[39,40] and a titanium dioxide enrichment kit (Thermo Fisher, https://www.thermofisher.cn/order/catalog/product/88303). In brief, the peptides were re-dissolved in 500 µL 50%ACN-1%TFA. Add 1.4 mg ZrAs-NPs to each tube to enrich phosphopeptides for 30 min. The beads were then separated by magnetic force and washed twice with 500 µL 50% ACN – 1% TFA. The enriched phosphopeptides were eluted twice with 500 µL 2.5% ammonia at room temperature for 15 min. The supernatant solution and washing solution after enrichment with ZrAs-NPs beads are subsequently enriched with a titanium dioxide enrichment kit. The phosphopeptides enriched by the two methods in different fractions were mixed together and dried in vacuum centrifuge for further use, separately.

NanoLC-MS/MS analysis

The NanoLC-MS/MS detection system consisted of an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific) coupled to an EASY-nLC 1200 nanoHPLC system equipped with an Acclaim™ PepMap™ 100 pre-column (20 mm × 75 µm, 3 µm) and a Acclaim™ PepMap™ RSLC C18 analytical column (150 mm × 75 µm, 2 µm). The mobile phases for the NanoLC separation were 0.1% FA in H_2O (A) and in 80% ACN/20% H_2O (B), respectively, at a flow rate of 300 nL/min. We run a 110-min gradient to separate these peptides (0-7 min, 2-7% B; 7-69 min, 7-20% B; 69-90 min, 20-35% B; 90-95 min, 35-95% B; 95-99 min, 95% B; 99-101 min, 95-50% B; 101-105 min, 50% B; 105-107 min, 50-2% B; 107-110 min, 2% B). Aliquot (1 µL) of each HPLC fractions containing 0.1% formic acid was loaded, and elution from the analytical column was directly infused to the mass spectrometer for MS/MS analysis. The ion source type was set as NSI and the detector type was set as Orbitrap. For a full mass spectrometry survey scan, the target value was 4×10^5 and the scan ranged from 350 to 1,800 m/z at a resolution of 120,000 and a maximum injection time of 50 ms. For the MS2 scan, a duty cycle of 3 s was set with the top-speed mode. Only spectra with a charge state of 2–5 were selected for fragmentation by collision induced dissociation with a normalized collision energy of 35%. The MS2 spectra were acquired in the orbitrap at a resolution of 15,000 with an AGC target of 5,000 and a maximum injection time of 22 ms. For the MS3 scan, the number of SPS precursor was set as 8 and the MS2 isolation window was set as 3 m/z. The target value was 5,000 and the scan ranged from 100 to 400 m/z at a resolution of 50,000 and a maximum injection time of 86 ms. Eight precursors were selected for fragmentation by higher-energy collision dissociation with a normalized collision energy of 65%.

or was set as 8 and the MS2 isolation window was set as 3 m/z. The target value was 5,000 and the scan ranged from 100 to 400 m/z at a resolution of 50,000 and a maximum injection time of 86 ms. Eight precursors were selected for fragmentation by higher-energy collision dissociation with a normalized collision energy of 65%.

r was set as 8 and the MS2 isolation window was set as 3 m/z. The target value was 5,000 and the scan ranged from 100 to 400 m/z at a resolution of 50,000 and a maximum injection time of 86 ms. Eight precursors were selected for fragmentation by higher-energy collision dissociation with a normalized collision energy of 65%.

The phosphopeptides were separated by a 120-min gradient (0-5 min, 2-7% B; 5-77 min, 7-20% B; 77-103 min, 20-35% B; 103-108 min, 35-95% B; 108-112 min, 95% B; 112-114 min, 95-50% B; 114-120 min, 50% B). The Orbitrap Fusion mass spectrometer was operated in the positive-ion mode at an ion transfer tube temperature of 320 °C. The positive-ion spray voltage was 2.3 kV. The full mass spectrometry survey scan resolution was set to 120,000 with an automated gain control (AGC) target of 4.1×10^5 for a scan range of 350–1,800 m/z and a maximum injection time of 110 ms. Only spectra with a charge state of 2–7 were selected for fragmentation by higher-energy collision dissociation with a normalized collision energy of 37%. The MS2 AGC target was set to 5,000 with a maximum injection time of 54 ms, and the dynamic exclusion was set to 20 s.

Proteomic/phosphoproteomic Data Analysis

Acquired MS/MS spectra were analyzed by Proteome Discovers 2.2 (Thermo Fisher Scientific) to identify and quantify proteins based on Protein Database (Version 10/25/2017, 42252 sequences) download from Uniport database. The taxonomy was Rattus norvegicus, and taxon identifier was 10116. Sequest HT 2.2 search engine was used for peptide matching. Trypsin was selected as the proteolytic enzyme, and four missed cleavages sites were allowed. The static modifications were carbamidomethylating at cystine, TMT labeling at N-terminus of all peptides. The dynamic modifications were oxidation at methionine, TMT labeling at lysine, methylation at lysine/glutamine/arginine, +Pt/+Pt(NH_3) at histidine/aspartic acid/glutamic acid/ methionine, and phosphorylation at serine/threonine/tyrosine. The peptide mass tolerance was set to 10 ppm. The fragment ion tolerance was set to 0.02 Da. The quantitative results were normalized based on the abundance of total peptides in each set. Abundance ratios were calculated for each protein identified in all groups other than controls by comparing with value of each protein to that of the same protein in control group. The false discovery rates of the peptide-spectrum matches (PSMs) and proteins were set to less than 1%. The other settings are the defaults of the PD software. The phospho-proteome data were searched against the same database with Proteome Discovers 2.2 with similar parameters, except for the phosphorylation sites probability was set as >75%.

Phosphopeptide Motif Analysis

Phosphopeptide sequences were submitted to MotifeR algorithm[31] (https://www.omicsolution.org/wukong/motifeR/) for the identification of known and novel phosphorylation motifs present in our data set. Only confidently identified phosphorylation sites with site probability >75% were used in analysis. Sequences were firstly aligned with the background database (9606-Homo sapiens) and truncated into uniform length sequences for the following analysis. Sequences were centered on each phosphorylation site and extended to 15 amino acids (±7 residues). When the site is located in the N/C-terminal of the protein, the sequence was filled up with up to 15 amino acids with the required number of “X”. The significance threshold was set to $P < 10^{-6}$ and the occurrence threshold was set as 20. This website can also generate logo-like representations to graphically display the identified motif. Moreover, according to the kinase-substrate information given in the literature, the upstream kinases of these significantly regulated phosphorylated peptides can also be given.

Bioinformatics Analysis

The proteins identified and quantified by MS/MS, of which false-discovery rates (FDRs) are <0.01, *p*-values <0.05 and abundance ratios (sample vs control) >1.2 or <0.833 (fold change >1.2 or <-1.2), were uploaded into the STRING (https://string-db.org) 10.0 data pool for Gene Ontology (GO) and protein-protein interaction (PPI) network analysis. The Uniport accession ID of the differentially expressed proteins (DEPs) in the DL/C group and L/C group were inputted into the list of gene name box, and Homo sapiens was selected for the analysis. Ingenuity Pathway Analysis (IPA, QIAGEN, 2021)[52] was used to analyze the core pathways, diseases, toxicities and functions associated with the DEPs. IPA analysis maps each protein to the corresponding molecule in the Ingenuity Pathway Knowledge Base, which is available at the Ingenuity System’s web site (http://www.ingenuity.com).

Statistical analysis

All data were analysed with Origin 8.0 and Microsoft offices excel 2010 software. Statistically significant differences were determined using unpaired t-tests, Student′s t-tests, Fisher's exact test right-tailed, as applicable and according to test requirements. A two-sided P value < 0.05 was considered statistically significant. * indicates $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

DATA AVAILABILITY

The raw sequencing data are deposited on Mendeley Data, and available on https://data.mendeley.com/datasets/k9sm56gj5y/1, https://data.mendeley.com/datasets/2tv7ntcgh6/1, https://data.mendeley.com/datasets/4zj3kz34c5/1, https://data.mendeley.com/datasets/dmyfkn7pnz/1, https://data.mendeley.com/datasets/dw3v4sm9nb/1, https://data.mendeley.com/datasets/xb6s7vzcwm/1, https://data.mendeley.com/datasets/75dyzmvbrz/1, https://data.mendeley.com/datasets/s7jbxvwf4z/1, and https://data.mendeley.com/datasets/fv6f9rpg5m/1.

The flow cytometry data contained in this article are deposited on FlowRespository with a Repository ID of FR-FCM-Z45H on https://flowrepository.org/experiments/4273.

The proteomics and phosphoproteomics data are deposited on Science Data Bank, and available on https://www.scidb.cn/en/anonymous/RmZhZUF6.

Acknowledgments

The authors thank the National Nature Science Foundation of China (No. 21575145, 22377130, and 21927804). Y.Z. was also supported by Beijing Natural Science Foundation, No. 2232034, National Key Research and Development Program of China, No. 2023YFF0713801. P.J.S. research on platinum is supported by Anglo American Platinum. The Medical Research Council, and the EPSRC (EP/P030572/1). The authors thank Professor Zhixin Tian from Tongji University for the help in analysing the number of phosphorylation sites.

AUTHOR CONTRIBUTIONS

F.Y.W. and Y.Z. conceived the project. L.Y.Q., Y.Z.H, Q.L., Y.X. and T.T.F. carried out the expression of HMGB1a and construction of HMGB1a-functionilizaed microprobes. L.Y.Q., B.B. X., Y.Y.Z., F.F.J., X.J.L., P.J.S., Y.Z. and F.Y.W. analysed sequencing data. L.Y.Q., W.C.Y. Y.Z. and X.K.L. performed bioinformatics analysis and Western Blot assays. L.Y.Q., F.Y.W. P.J.S., Y.Z. and S.J.W. wrote and revised the manuscript with inputs from all authors.

CONFLICT OF INTEREST

The authors have no conflict to declare.

Extended Data Table 1 – 11 are separately provided in Excel format.