Transcriptome profiling reveals Silibinin dose-dependent response network in non-small lung cancer cells

Jagan Mohan Kaipa\textsuperscript{1,2,3}, Vytaute Starkuviene\textsuperscript{2,4}, Holger Erfle\textsuperscript{2}, Roland Eils\textsuperscript{5,6} and Evgeny Gladilin\textsuperscript{2,7,8}

\textsuperscript{1}Helmholtz Center for Infection Research, Braunschweig, Germany
\textsuperscript{2}BioQuant, University Heidelberg, Heidelberg, Germany
\textsuperscript{3}Theoretical Bioinformatics, German Cancer Research Center, Heidelberg, Germany
\textsuperscript{4}Institute of Biosciences, Vilnius University Life Science Center, Vilnius, Lithuania
\textsuperscript{5}Center for Digital Health, Berlin Institute of Health and Charité Universitätsmedizin Berlin, Berlin, Germany
\textsuperscript{6}Health Data Science Unit, Heidelberg University Hospital, Heidelberg, Germany
\textsuperscript{7}Leibniz Institute of Plant Genetics and Crop Plant Research, Seeland, Germany
\textsuperscript{8}Applied Bioinformatics, German Cancer Research Center, Heidelberg, Germany

\textbf{ABSTRACT}

Silibinin (SIL), a natural flavonolignan from the milk thistle (Silybum marianum), is known to exhibit remarkable hepatoprotective, antineoplastic and EMT inhibiting effects in different cancer cells by targeting multiple molecular targets and pathways. However, the predominant majority of previous studies investigated effects of this phytocompound in a one particular cell line. Here, we carry out a systematic analysis of dose-dependent viability response to SIL in five non-small cell lung cancer (NSCLC) lines that gradually differ with respect to their intrinsic EMT stage. By correlating gene expression profiles of NSCLC cell lines with the pattern of their SIL IC50 response, a group of cell cycle, survival and stress responsive genes, including some prominent targets of STAT3 (BIRC5, FOXM1, BRCA1), was identified. The relevancy of these computationally selected genes to SIL viability response of NSCLC cells was confirmed by the transient knockdown test. In contrast to other EMT-inhibiting compounds, no correlation between the SIL IC50 and the intrinsic EMT stage of NSCLC cells was observed. Our experimental results show that SIL viability response of differently constituted NSCLC cells is linked to a subnetwork of tightly interconnected genes whose transcriptomic pattern can be used as a benchmark for assessment of individual SIL sensitivity instead of the conventional EMT signature. Insights gained in this study pave the way for optimization of customized adjuvant therapy of malignancies using Silibinin.

\textbf{INTRODUCTION}

The ability of malignant cells to attain drug resistance and to escape cell death frequently observed in different types of cancer represents the major challenge to chemical tumor therapy. Several evolutionary conserved mechanisms mediate elevated survival capabilities
and drug resistance of cancer cells including inhibition of apoptotic pathways, alteration and enhancement of metabolism, DNA damage repair, drug inactivation or alteration, genetic and epigenetic activation of stress response and proliferation programs such as epithelial-mesenchymal transition (EMT) (Michael & Doherty, 2005; Hang, Cai & Fan, 2013; Housman et al., 2014; Hanahan & Weinberg, 2011). As a consequence of elevated environmental adaptation and self-reprogramming capabilities, cancer cells often attain resistance against targeted drugs and even drug combinations by bypassing affected pathways (Patel & Rothenberg, 1994; Clarke et al., 2019). In most of the tumors, multiple survival mechanisms and pathways are active in parallel. A network of dozens tightly interconnected genes rather than just few linear signaling pathways maintain abnormal survival and resistance capabilities of cancer cells (Gladilin & Eils, 2017).

A promising approach to overcome the limitations of conventional therapy is a combination of targeted and sensitizing drugs (Chandarlapaty et al., 2010; Al-Lazikani, Banerji & Workman, 2012). In view of often unavoidable side effects of targeted therapy, sensitizing compounds with a low toxicity for healthy cells are preferred that are still capable to weaken abnormally functioning tumor cells. There is an increasing body of evidence that secondary plant metabolites such as various polyphenolic compounds exhibit distinctive antineoplastic properties that make them promising sensitizing agents for combined tumor therapy (Gerhäuser, 2012; Wang et al., 2016; Dayem et al., 2016; Thomford et al., 2018). Some of these natural compounds have a long history of usage as nutrition supplements in traditional medicine, and were empirically proven to be well tolerated.

One of the antineoplastic phytocompounds increasingly gaining attention in the last two decades is Silibinin (also known as Silybin or Silymarin)—a flavonolignan from the milk thistle Silybum marianum. Silibinin is a mixture of two diastereoisomers Silybin A and Silybin B, at a ratio of 1:1 (Agarwal et al., 2003). Originally known as a hepatoprotective dietary drug (Machicao & Sonnenbichler, 1977; Sonnenbichler et al., 1999; Flora et al., 1998; Vargas-Mendoza, 2014; Hellerbrand et al., 2016), in recent years Silibinin has been shown to exhibit remarkable antineoplastic properties crossover different types of tumors which was attributed to different molecular mechanisms and signaling pathways (Fraschini, Demartini & Esposti, 2002; Gazak, Walterova & Kren, 2007; Ramasamy & Agarwal, 2008; Ting, Deep & Agarwal, 2013; Polachi et al., 2016). Meanwhile more than 1,500 studies on hepatoprotective and antineoplastic effects of Silibinin and their mechanisms in different tissues and cells were published (PubMed, 2019). The predominant majority of these works were, however, performed with one particular cell line and typically focused on a few molecular targets and signaling/metabolic pathways directly or indirectly affected by Silibinin, see Table 1. Primary molecular targets and mechanisms of SIL action in particular type of cancer tissue/cells were usually hypothesized and rarely investigated mechanistically. As direct molecular binding targets of Silibinin several drivers and mediators of malignant transformation were reported, see Table 2.

As a multitarget compound with a broad spectrum of antineoplastic action Silibinin combines the ability to selectively reduce viability of cancer cells in a dose dependent manner (Katiyar, Roy & Baliga, 2005; Mokhtari, Motamed & Shokrgozar, 2008; Zhan et al., 2011; Su et al., 2013; Wang et al., 2014b) with low toxicity for normal tissues even in the range of
<table>
<thead>
<tr>
<th>Cancer tissue, cell type</th>
<th>Molecular mechanisms</th>
<th>Lit.</th>
</tr>
</thead>
<tbody>
<tr>
<td>H. lymphoma, U-937</td>
<td>Inh: TNF, NF-κB, MAPK8</td>
<td>Manna et al. (1999)</td>
</tr>
<tr>
<td>M. SKH-1 epidermis</td>
<td>Up: p53, p21 Down: CCNE1, CCND1, CDC25C, CCNB1 Inh: CDK2, CDK4, CDC2</td>
<td>Dhanalakshmi et al. (2004)</td>
</tr>
<tr>
<td>H. colorectal cancer, SW480</td>
<td>Inh: PIK3CA-Akt-mTOR Act: MAP2K1/2-MAPK1/3</td>
<td>Raina et al. (2013)</td>
</tr>
<tr>
<td>H. breast cancer, MCF-7</td>
<td>Inh: HSP90 Act: MAP2K1/2-MAPK1/3</td>
<td>Zhao et al. (2011)</td>
</tr>
<tr>
<td>H. lung cancer, A549</td>
<td>Inh: PI3K-Akt-MAPK</td>
<td>Chen et al. (2005)</td>
</tr>
<tr>
<td>H. breast cancer, MCF-7, MDA-MB-231</td>
<td>Inh: Notch-1 Down: ERK, Akt, AIF, CASP3</td>
<td>Kim et al. (2014b)</td>
</tr>
<tr>
<td>H. colorectal adenocarcinoma, LoVo</td>
<td>Inh: GLUT1</td>
<td>Catanzaro et al. (2018)</td>
</tr>
<tr>
<td>H. glioblastoma, A172, SR</td>
<td>Inh: mTor, Yap</td>
<td>Bai et al. (2018)</td>
</tr>
<tr>
<td>H. prostate cancer, DU145, PC3</td>
<td>Up: p27, p21 Down: CDK4, CDK6, CDK2, CCNE1, CCND1, CCNB1</td>
<td>Davis-Searles et al. (2005), Mokhtari, Motamed &amp; Shokrgozar (2008)</td>
</tr>
<tr>
<td>H. lung cancer, PC-9</td>
<td>Up: CDH1</td>
<td>Cufí et al. (2013)</td>
</tr>
<tr>
<td>H. bladder cancer, T24</td>
<td>Inh: CTNNB1/ZEβ1</td>
<td>Wu et al. (2013)</td>
</tr>
<tr>
<td>H. breast cancer, MDA-MB-231</td>
<td>Inh: CXCR4</td>
<td>Wang et al. (2014a)</td>
</tr>
<tr>
<td>H. breast cancer, MDA-MB-231</td>
<td>Down: CDC42, D4-GDI</td>
<td>Dastpeyman et al. (2012)</td>
</tr>
<tr>
<td>H. leukemia, THP-1</td>
<td>Inh: p65, ICAM-1</td>
<td>Chen et al. (2014)</td>
</tr>
<tr>
<td>H. lung cancer, A549</td>
<td>Inh: STAT1, STAT3, NF-κB</td>
<td>Chittezhath et al. (2008)</td>
</tr>
<tr>
<td>H. hepatoma, Hep3B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H. prostate cancer, PCA</td>
<td>Down: SREBP1/2 Up: AMPK</td>
<td>Nambiar et al. (2014)</td>
</tr>
</tbody>
</table>

(continued on next page)
### Table 1 (continued)

<table>
<thead>
<tr>
<th>Cancer tissue, cell type</th>
<th>Molecular mechanisms</th>
<th>Lit.</th>
</tr>
</thead>
<tbody>
<tr>
<td>H. colon cancer</td>
<td>Down: CCND1, CDK4, CDK6</td>
<td>Kauntz et al. (2012)</td>
</tr>
<tr>
<td></td>
<td>Up: Bax</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Down: Bcl-2, IL1β, TNFα, MMP7</td>
<td></td>
</tr>
<tr>
<td>H. colon cancer, SW480</td>
<td>Act: CASP3, CASP8</td>
<td>Kauntz (2012)</td>
</tr>
<tr>
<td>H. lung cancer</td>
<td>Down: HIF1A, TNFα, NK-kB, STAT3</td>
<td>Tyagi et al. (2009)</td>
</tr>
<tr>
<td></td>
<td>Up: Ang-2, Tie-2, TIMP-1, TIMP-2</td>
<td></td>
</tr>
<tr>
<td>H. cervical cancer, HeLa</td>
<td>Up: ROS, NOS</td>
<td>Fan et al. (2011)</td>
</tr>
<tr>
<td>H. lung cancer, MCF-7</td>
<td>Down: Erα, mTor, ERK</td>
<td>Zheng et al. (2015)</td>
</tr>
<tr>
<td>H. glioma, U87, U251</td>
<td>Down: PI3K, FOXM1</td>
<td>Zhang et al. (2015)</td>
</tr>
<tr>
<td>H. breast cancer, MCF7</td>
<td>Down: Bcl-2, BRCA1</td>
<td>Pirouzpanah et al. (2015)</td>
</tr>
<tr>
<td></td>
<td>Down: Nanog, CTNNB1, CDKN2A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Up: CDH1</td>
<td></td>
</tr>
<tr>
<td>H. prostate cancer, DU145</td>
<td>Inh: STAT3, pSTAT3</td>
<td>Agarwal et al. (2007)</td>
</tr>
<tr>
<td>H. gastric cancer, MGC803</td>
<td>Inh: STAT3, pSTAT3</td>
<td>Wang et al. (2014a)</td>
</tr>
<tr>
<td>H. breast cancer, MDA-MB468, BT20</td>
<td>Inh: STAT3</td>
<td>Kim et al. (2014a)</td>
</tr>
</tbody>
</table>

**Notes.**
Abbreviations: Up, upregulation; Down, downregulation; Inh, inhibition; Act, activation; H, human; M, mice.

### Table 2 Overview of reported direct molecular binding targets of Silibinin.

<table>
<thead>
<tr>
<th>Protein name</th>
<th>Gene symbol</th>
<th>Lit.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CXC-Motiv-Chemokinrezeptor 4</td>
<td>CXCR4</td>
<td>Wang et al. (2014a)</td>
</tr>
<tr>
<td>Mitogen-Activated Protein Kinase 11</td>
<td>MAPK11</td>
<td>Youn et al. (2013)</td>
</tr>
<tr>
<td>G-protein coupled receptor 12</td>
<td>GPR120</td>
<td>Chinthakunta et al. (2018)</td>
</tr>
<tr>
<td>Cyclooxygenase-2</td>
<td>COX-2</td>
<td>Malik, Manan &amp; Kirza (2017)</td>
</tr>
<tr>
<td>Phospholipase A2</td>
<td>PLA2</td>
<td>Malik, Manan &amp; Kirza (2017)</td>
</tr>
<tr>
<td>Aldo-Keto Reductase</td>
<td>AKR1D1</td>
<td>Malik, Manan &amp; Kirza (2017)</td>
</tr>
<tr>
<td>Family 1 Member D1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Core 1 β1,3-galactosyltransferase</td>
<td>C1GALT1</td>
<td>Lin et al. (2018)</td>
</tr>
<tr>
<td>β-catenin</td>
<td>CTNB1</td>
<td>Ifitikhar &amp; Rashid (2014)</td>
</tr>
<tr>
<td>Epidermal Growth Factor Receptor</td>
<td>EGFR</td>
<td>Hosen et al. (2016)</td>
</tr>
<tr>
<td>Heat Shock Protein 90</td>
<td>HSP90</td>
<td>Zhao et al. (2011), Riebold et al. (2015), Cys et al. (2019)</td>
</tr>
</tbody>
</table>
high, therapeutically relevant doses of more than 1,500 mg/day (Soleimani et al., 2019). However, the knowledge of primary targets alone does not provide yet a reliable criterion for quantitative assessment of SIL efficiency in application to a particular cancer tissue/cell. One of the global factors known to be relevant to the drug dose response is the intrinsic EMT stage of cancer cells. Accordingly, cells exhibiting mesenchymal pheno-/genotypic profiles are more drug-resistant to antineoplastic drugs than epithelial cells (Tan et al., 2014). It is, however, not known whether this rule also applies to Silibinin. Here, we address this question by a systematic analysis of dose-dependent viability response to SIL in five non-small cell lung cancer (NSCLC) lines (H1650, H1975, A549, H838, H2030) that gradually differ with respect to their intrinsic EMT stage defined by the transcriptomic signature from (Byers et al., 2013). The relationship between SIL IC50 viability response and gene expression profiles of five NSCLC cell lines was compared with a reference compound, Withaferin-A, which is known to exhibit dose-dependent EMT-inhibiting effects (Vanden Berghe et al., 2012; Vyas & Singh, 2014; Gladilin, Gonzalez & Eils, 2014). Our experimental results show that in contrast to WFA SIL does not exhibit a EMT-conform dose–response. Instead, SIL viability response of NSCLC cells turns out to correlate with the expression level of cell cycle, survival and stress responsive genes including some prominent targets of STAT3.

MATERIALS AND METHODS

Cell culturing

Human lung adenocarcinoma cell lines H1650, H1975, A549, H838, H2030 were obtained from ATCC and cultured in DMEM (Dulbecco’s Modified Eagle Medium) medium supplemented with 10% fetal calf serum (FCS) and 100 U/ml penicillin G and 100 µg/ml streptomycin sulfate at 37 °C in a humidified 5% CO2 incubator. To ensure the ample number of cell count, cells were cultured three days prior to cell seeding. When the culture plates became confluent, NSCLC cells were detached and collected in a sterile Falcon tube with a pre-warmed medium for transfection. Cells were subjected to centrifugation (at 1000 rpm and room temperature for 5 min) and the supernatant was aspirated. The cells were resuspended in the medium for transfection. The number of cells/ml was determined by using Neubauer counting chamber and cells were diluted to get a concentration of 12000 cells/100 µl of the medium. After careful optimization of transfection efficiency, 12000 cells/well in 100 µl was found to show good transfection rate.

Silibinin and Withaferin-A compounds

Silibinin (C25H22O10, mol. weight 482.44) and Withaferin-A (C28H38O6, mol. weight 470.60) both were purchased from Sigma-Aldrich (Germany). Both compounds were dissolved in dimethylsulfoxide (DMSO; Sigma-Aldrich) to make a stock solution. The final concentration of DMSO in the culture medium did not exceed 0.5% which has no detectable effects on cells.

CellTiter-Blue cell viability assay

Cells seeded into 96-well plates in sextuplicate at a density of 15 × 10³ cells per well. Twenty-four hours after seeding, the cells were treated with various concentrations of
Withaferin-A and Silibinin for totally 24 h. To measure the viability of cells, CellTiter-Blue® Viability Assay (Promega, G8081) was applied according to the manufacturer’s instructions. Incubation with the dye for 60 min was followed by measurement of the fluorescence with the infinite F200 pro Reader (TECAN). A blank well without cells was measured to determine the background. After subtraction of the background, cell dilution series enabled due to a direct correlation of the signal intensities with the cell number, the verification of absolute cell numbers.

**Determination of IC50 from CTB measurements**

The half maximal inhibitory concentration (IC50) was determined from series of dose-dependent CTB measurements using following basic steps. First, the raw CTB intensity measurements \( I_i \) were normalized by the intensity level \( I_0 \) of the reference (untreated) probe:

\[
nI_i = 100 \frac{I_i}{I_0}.
\]

(1)

Subsequently, the average pattern of normalized dose response \( anI_i \) from \( N \) different measurements was calculated:

\[
anI_i = \frac{1}{N} \sum_{j=1}^{N} nI_i(j).
\]

(2)

Finally, the IC50 value of a particular cell line and measurement condition (e.g., duration of treatment) was determined from the average normalized pattern of IC50 dose response by fitting the Hill’s sigmoid function \(Spiess & Neumeyer, 2010; Sebaugh, 2011\)

\[
anI(D) = \min + \frac{\max - \min}{1 + \left(\frac{D}{IC50}\right)^H},
\]

(3)

where \( D \) is the drug dose, \( \min \) and \( \max \) are the minimum and maximum values of \( anI \), \( IC50 \) and \( H \) are the IC50 and the Hill’s coefficient values that are determined from the fit. The nonlinear least square fit of the Hill’s equations was performed automatically using the MATLAB R2019b (The Mathworks, Inc.). To characterize relative differences in viability response \( V_i \) among five NSCLC cell lines \( (i = 1..5) \), normalized IC50 values \( nV_i \)

\[
nV_i = \frac{V_i - V_{\min}}{V_{\max} - V_{\min}} \in [0, 1]
\]

(4)

and their ranks \( R_i \)

\[
R_i = \text{rank}(V_i, [V_1, V_2, V_3, V_4, V_5]).
\]

(5)

were calculated.

**Determination of total cell count using quantitative image analysis**

Quantitative image analysis was performed to acquire the total cell count. Immediately after the measurement of the fluorescence signal, the medium is aspirated from 96well plate containing NSCLC cells and washed with PBS solution twice to ensure the complete removal of dead cells. Since dead cells do not attach to the surface and freely float in the
medium, washing twice with PBS removes approximately all the dead cells from the well. After washing, 70 µl cell fixation solution (PFA 3% + Hoechst dye) was added to each well for the fixation and nuclei staining of cells by a single step. Cells were incubated at 4 °C for 48 h. Wells were washed with PBS twice after incubation to remove unbound Hoechst dye and analyzed under a microscope (Olympus FV 3000) for counting the stained nuclei. Finally, the cell counting was performed manually with the help of images obtained from the microscope.

**In vitro cell migration assay**

For measurement of cell motility, 2D cell migration assay from (Marwitz et al., 2016) was used. Cells were seeded in 24-well plate (Zell Kontakt 3231-20) at a density of 2,000–5,000 cells/well (depends from cell lines). Twenty-four hours after seeding the cells were treated with various concentrations of Withaferin-A and Silibinin for a total of 24 h and after stained with Hoechst (Sigma) for 1 h. Images were acquired in 20 min intervals for 48 h using environment-controlled microscope (IX81, Olympus) equipped with an UPlanSApo 10 × /0.4 objective lens (Olympus). Nine positions per well (3 × 3 grid) were imaged and stitched with a ImageJ plugin (Preibisch, Saalfeld & Tomancak, 2009). Single cell tracking was performed with ImageJ Mtrack2 plugin. Speed of each tracked cell was calculated by dividing total travelled distance by total time, for which cell was tracked. Persistence was calculated by dividing the distance between the first and the last point, where the cell was tracked, by total travelled distance. Resulting number was multiplied by the square root of time, for which cell was tracked divided by maximal possible tracking time, in order to penalize cells, which were tracked for a shorter period of time.

**Immunofluorescence and microscopic imaging**

Cells were fixed for 5 min in 3.7% PFA in PBS at RT, permeabilized with ice-cold 0.1% TritonX-100 in PBS for 5 min, blocked for 30 min with Blocking solution (3% BSA, 2.5% FCS) and stained with primary antibodies to vimentin and actin (Abcam, Cambridge, UK) for 1 h. After staying cells were washed 3 times in PBS and stained with secondary antibodies, Phalloidin-BODIBY (Thermo Fisher) and Hoechst (Sigma) for 30 min. Imaging was performed using the confocal laser scanning microscope (Olympus FV 3000) with 64x oil immersion lens to visualize the actin and vimentin network as well as the nucleus in the cells.

**Gene expression data**

Gene expression (GE) data of H1650, H1975, A549, H838, H2030 cell lines from GSE47206 (El-Chaar et al., 2014) were used for computational analysis of correlation between transcriptomic and IC50 profiles of NSCLC cells. Analysis was performed only for transcripts with the detection significance of \( p < 0.05 \) (P,M) in all five NSCLC cell lines. Transcripts with the non-significant level of detection (A) were excluded from analysis.

**GO enrichment analysis**

Gene ontology enrichment analysis was performed using STRING v11 (Szklarczyk et al., 2019) with the significance level of \( p < 0.05 \) after multiple testing correction (Benyamini & Hochberg, 1995) for each functional classification framework (GO, KEGG, InterPro, etc.).
<table>
<thead>
<tr>
<th>Gene</th>
<th>Full Name</th>
<th>Sense siRNA Sequence</th>
<th>Antisense siRNA Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIRC5</td>
<td>Baculoviral IAP repeat containing 5</td>
<td>GGACCACCGCAUCUCUACATT</td>
<td>UGUAGAUGCGGGUGGUCCTT</td>
</tr>
<tr>
<td>BIRC5</td>
<td>Baculoviral IAP repeat containing 5</td>
<td>GCAGGUUCUUAUCUGUCATT</td>
<td>UGACAGAUAGGAACUGCGAG</td>
</tr>
<tr>
<td>BIRC5</td>
<td>Baculoviral IAP repeat containing 5</td>
<td>CAAAAGAAACAAUAAUAATT</td>
<td>UUAUUGUUGGUCUUCUUGCA</td>
</tr>
<tr>
<td>FOXM1</td>
<td>Forkhead box M1</td>
<td>GCUCUAACUGGUACCUAUTC</td>
<td>AUAGGUACCCAGGAUGAGCTG</td>
</tr>
<tr>
<td>FOXM1</td>
<td>Forkhead box M1</td>
<td>CACUAUCAACAAUUGACCUATT</td>
<td>UAGGUACCCAGGAUGAGCA</td>
</tr>
<tr>
<td>FOXM1</td>
<td>Forkhead box M1</td>
<td>GGAUCAAGAUAAUUAACATT</td>
<td>UGGUAAUAAUACUGUACCCA</td>
</tr>
<tr>
<td>BRCA1</td>
<td>Breast cancer 1, early onset</td>
<td>GGGGAUACCAUGCAACUAATT</td>
<td>UUAUGUUGCAUGGAUACCCCTC</td>
</tr>
<tr>
<td>BRCA1</td>
<td>Breast cancer 1, early onset</td>
<td>CAGCUACCCUUUCCCAUAATT</td>
<td>UAGAUGGAAAGGUGACUGTT</td>
</tr>
<tr>
<td>BRCA1</td>
<td>Breast cancer 1, early onset</td>
<td>CAUGCAACAAUACCUGAUATT</td>
<td>UACGAGUAAUUGUGCAUGGT</td>
</tr>
</tbody>
</table>

**Transfection of NSCLC cells**

All siRNA used for knockdown of BIRC5, FOXM1, BRCA1 was purchased from Ambion® ThermoFisher Scientific, see Table 3.

Transfection experiments were carried out by lipofection technique using Lipofectamine® 2000 (Invitrogen), with transient expression. Two types of transfection methods were performed in the optimization of transfection efficiency: reverse and forward transfection.

**Forward and reverse transfection method**

Forward transfection is a conventional method of transfection where cells are first seeded for 24 h and stored at 37 °C and 5% CO₂. After incubation, cells are transfected as per company’s (Invitrogen) protocol. In reverse transfection, unlike conventional transfection, the genetic material is coated on the bottom of well plates prior to the seeding of NSCLC cells into multi-well plate. As the order of adding genetic material is reverse compared to the conventional method it is named as reverse transfection method. Multi-well plates are coated with required siRNA to be transfected along with gelatin, fibronectin and transfection reagent (Lipofectamine® 2000) stored at 4 °C for 28 h and stored in dry place. Later, NSCLC cells were seeded into these wells, incubated for 28 h to attain transfected cells.

**Detection of siRNA mediated gene silencing**

Estimation of transfection efficiency of the NSCLC cells was performed by blocking PLK1 (polo-like-kinase-1) enzyme using RNA silencing technique via small interfering RNA (siRNA). Expression of PLK1 has several functions like the initiation of mitosis, cytoplasmic separation, mitotic spindle formation and membrane formation in telophase of mitosis by phosphorylating mitotic kinesin-like protein-1. Silencing of PLK1 leads to mitotic cell arrest and the centrosomes lose the ability to nucleate microtubules. This effect can be observed under a microscope with star shape nuclear phenotype. Presence of star shape nucleus indicates that the PLK1 was successfully silenced that in turn confirms that the cells are transfectable.
**Immunoblotting**

The confirmation of knockdown efficiency was carried out using Western Blot analysis. Harvesting and samples for analysis were prepared by removing supernatant from the cells followed by addition of 500 \( \mu l \) of PBS. Cells were resuspended and transferred to 1.5 ml for centrifugation for 5 min at 1,600 RPM. After centrifugation, PBS was removed without disturbing cell pellet. 25 \( \mu l \) of 1x Laemli buffer containing DTT (final concentration 100 mM) was added and incubated for 10 min at 95 °C. Proteins in the sample were separated by SDS-PAGE in a 10% resolving gel and transferred onto methanol-activated PVDF membranes by semi-dry blotting technique using transfer buffer at a constant current of 130 V until the sample reaches the resolving gel. As soon as the sample reaches resolving gel electrophoresis is run at 90V. The membrane was blocked with 5% milk powder, 0.5% BSA in PBST for 1 h at room temperature. After blocking, the membrane was incubated with specific primary antibody in 5% (w/v) milk powder/PBST overnight at 4 °C. To remove the unbound primary antibody three subsequent washings for 10 min with PBST were performed. Then the membrane was incubated with the horseradish peroxidase (HRP) coupled secondary antibody in 5% (w/v) milk powder, 0.5% BSA in PBST for 1 h at room temperature. Three subsequent PBST washing were performed to ensure complete removal of the secondary antibody. Detection of proteins was carried out by using ECL chemiluminescent immunodetection system with different exposure time using INTAS Fluoreszenz u. ECL Imager. Subsequently, quantitative analysis of Western Blot images was performed using ImageJ according to a standard protocol (http://www.openwetware.org/wiki/Protein_Quantification_Using_ImageJ). Figure S5 shows exemplary analysis of transient knockdown of target proteins FOXM1 and BIRC5 in A549 and H838 NSCLC cells using small-interfering RNA (siRNA) vs \( \beta \)-actin as a loading control.

**RESULTS**

The intrinsic EMT stage of NSCLC cell lines was ranked according to the expression level of E-cadherin \((CDH1)\) as suggested by Byers et al. (2013), see Table S1. Accordingly, five NSCLC cell lines (H1650, H1975, A549, H838, H2030) exhibit gradual difference in their intrinsic EMT stage: from most epithelial (H1650) to most mesenchymal (H2030), respectively.

Dose-dependent response of all five NSCLC cell lines to 24 h, 48 h and 72 h exposure with Silibinin (SIL) and the reference compound, Withaferin-A (WFA), measured via the CellTiter-Blue \(^ \text{R} \) Cell Viability Assay was used to determine IC50 values of both compounds. Figure 1 shows dose response curves and SIL/WFA IC50 values of all five NSCLC cell lines.

As one can see from the overview of SIL/WFA IC50 ranks of all five NSCLC cell lines in Fig. 2, the pattern of WFA IC50 values exhibits a strong correlation with the intrinsic EMT stage of NSCLC cells, while SIL IC50 does not show such a correlation. Strikingly, all cell lines with exception of H838 have a comparable level of SIL IC50.

To test possible dependency of the drug response on the intrinsic EMT stage of NSCLC cells, Spearman (rank) correlation between the pattern of normalized SIL and WFA IC50
Figure 1  Dose-dependent 24 h, 48 h, 72 h response of NSCLC cells lines to Silibinin (A–F) and Withaferin-A (G–L): (A, G) H1650, (B, H) H1975, (C, I) A549, (D, J) H838, (E, K) H2030, (F, L) mean SIL IC50 values for all five NSCLC cell lines. Error bars indicate stdev of measurements performed with three replicates.
and the EMT 58 gene signature from Byers et al. (2013) was calculated. From analysis of correlations between expression of 58 EMT genes and the pattern of SIL/WFA IC50 response, it follows that the WFA IC50 of NSCLC cells exhibits a significantly higher correlation with the EMT stage than the SIL IC50 response, see Table S1. Examples of significant WFA and poor SIL correlations with two major biomarkers of EMT (e.g., CDH1 and ZEB1) are shown in Fig. 3. While gradient of WFA IC50 values from epithelial to mesenchymal cell lines is consistent with previous observations of elevated drug-resistance and survival potential of mesenchymal cells (Tan et al., 2014), low correlation of SIL IC50 with EMT gene expression (GE) appear to contradict this prevailing view.

In order to dissect further possible genes correlating with the pattern of SIL IC50 among five NSCLC cell lines, correlation analysis was performed at the whole genome scale under consideration of reliability of transcriptomic signals (i.e., M/P accepted, A excluded), see Table S2. From this analysis 144 genes with a positive correlation and significance level of \( p < 0.05 \) were identified. Gene ontology (GO) analysis of 144 genes positively correlating with the SIL IC50 signature of five NSCLC cell lines shows a significant enrichment in GO categories related to cell cycle, G2/M transition, nuclear localization, DNA-replication and repair, and related signaling and metabolic processes and functions, see Table S2. Furthermore, a statistically significant overlap between this group of 144 genes and 90 ‘high-communicability’ genes from our previous pan-cancer analysis study (Gladilin & Eils, 2017) was detected (12/90=13.3 %, \( p < 0.001 \) hypergeometric test: \( hgt(22277, 144, 90, 12) = 5.7e^{-13} \)), see Table S3. Notably, this set of 12 cell cycle process related genes includes three genes that are known to be prominent targets of STAT3 transcription factor: BIRC5 (Alvarez & Frank, 2004; Gritsko et al., 2006; Carpenter & Lo, 2014), FOXM1 (Mencalha et al., 2012), BRCA1 (Snyder, Huang & Zhang, 2007), see Fig. 4. Figure 5 shows correlation...
between gene expression patterns of these three STAT3 target genes and SIL IC50 response of five NSCLC cells vs. another three non-significantly correlating genes.

To evaluate the relevancy of computationally selected genes to SIL viability response of NSCLC cells, transient knockdown of BIRC5, FOXM1, BRCA1 genes was performed. Even though the reverse transfection method was considered to be more efficient providing higher transfection rates with minimal nucleic acid usage (Erfle et al., 2007), here we found out forward transfection to show a better performance crossover five tested NSCLC cell lines, see Fig. S1. Reverse and forward transfection of siRNA targeting PLK1 resulted in the comparable phenotype frequencies in all tested cell lines. A459 and H838 responded strongly in both methods, whereas, H1650 and H2030 showed no phenotype. This may be attributed either to difficulties in transfecting these cells or non-responsiveness to PLK1 depletion. Nearly 50% of H1975 cells showed prometaphase arrest when using forward transfection, but hardly any effect was observed under the conditions of reverse transfection. Consequently, the forward transfection protocol was used for transfection of H1975, A549, H838 cells with target siRNA.

Comparative viability measurements of H1975, A549, H838 cells with transiently knocked down BIRC5, FOXM1, BRCA1 genes confirmed their relevancy to SIL response with exception of BIRC5 knockdown in H838 cells, see Fig. 6. We draw the negative result
of BIRC5 knockdown in H838 cells back to constitutively low level of BIRC5 in H838 in comparison to H1975, A549 cell lines. The relative changes in SIL response between control and knocked down NSCLC cells observed in this work are in the order of magnitude comparable with previous observations for other small weight compounds and NSCLC cell lines (Gao et al., 2020).

**DISCUSSION**

There is abundant evidence that Silibinin exhibits remarkable antineoplastic effects crossover different cancer tissue/cell types by affecting multiple molecular targets and pathways. However, not much is known about factors responsible for individual SIL viability response of cancer cells. Here, we systematically analyzed the relationship between the patterns of SIL IC50 and gene expression in five NSCLC cell lines exhibiting gradual difference with respect to their intrinsic EMT stage and compared it with the reference compound Withaferin-A. Our experimental results showed that, differently from WFA and other drugs, sensitivity of NSCLC cells to Silibinin does not correlate with the intrinsic EMT stage. Instead, a subset of cell cycle, survival and stress responsive genes including three prominent targets of STAT3 (i.e., BIRC5, FOXM1, BRCA1) was found to exhibit significant correlation with the pattern of SIL IC50 in five NSCLC cell lines, see Fig. 7A. Subsequent evaluation of SIL viability response of transiently BIRC5, FOXM1, BRCA1 silenced NSCLC cells confirmed computationally predicted dependency of SIL dose on the expression level of these genes. Since the expression of STAT3 transcriptional targets but not STAT3 itself
significantly correlates with SIL IC50 of five NSCLC cell lines, we conclude that external and/or constitutive activation of STAT3, for example, due to gain-of-function mutations or enhanced upstream signaling (e.g., tyrosine kinase receptors, G protein coupled receptors, toll-like receptors, growth factor receptors) is responsible for differences in SIL sensitivity between different NSCLC cell lines, see Fig. 7B. In this point, our study confirms previous findings that inhibition STAT3 and its downstream targets is one of the major mechanisms of Silibin antineoplastic action. Furthermore, our observations confirm previous findings that expression of known drug targets does not necessarily exhibit a significant correlation with IC50 (Parca et al., 2019), and that downstream targets of those drug targets may instead represent a more reliable reference for IC50 correlation studies.

Our correlation analysis between viability and whole transcriptome profiles of five NSCLC cell lines provides a rich source of information for further investigation of anti-neoplastic and EMT-inhibiting effects of Silibinin in cancer cells. Among the group of 144 proteins whose expression positively correlates with the SIL IC50 there are also significantly enriched ontological categories related to cytoskeleton regulation and organization, see Table S2. Our previous studies (Gladilin, 2015) indicated remarkable inhibiting effects of Silibinin on cancer cell migration and overall reorganization of vimentin and actin networks (Supplementary Figs. S2 and S3) that can be attributed to these cytoskeleton-related genes, see Fig. S4. While GO enrichment in this work was performed using conventional statistical testing, rapidly evolving deep learning techniques (e.g., Cao et al., 2018; Le et al., 2019) hold the promise of proving more in-depth insights into large gene expression and cell viability data in the future.

Figure 5  Examples of genes with significant (FOXM1 (A), BIRC5 (B), BRCA1 (C)) and non-significant (TCF12 (D), FBXL11 (E), ANAPC13 (F)) correlation between the patterns of gene expression and normalized SIL IC50 in five NSCLC cell lines.  

Full-size DOI: 10.7717/peerj.10373/fig-5
Figure 6 Summary of cell viability measurements of BRCA1 (A–C), BIRC5 (D–F) and FOXM1 (G–I) knocked down NSCLC cell lines with and without Silibinin treatment of H1975 (A, D, G), A549 (B, E, H) and H838 (C, F, I) NSCLC cell lines (see Table S4 for raw data for these plots). Error bars indicate stdev of measurements performed with three replicates. Negative control = NSCLC cells + lipofectamine® 2000 + scramble siRNA, knockdown = NSCLC cells + lipofectamine® 2000 + siRNA. NS, \( p > 0.05 \) (non-significant); *, \( p < 0.05 \); **, \( p < 0.01 \); ***, \( p < 0.001 \).

Furthermore, our experimental results suggest that the intrinsic EMT stage is not a suitable biomarker of SIL dose response of cancer cells. Instead, the subset of SIL IC50 correlating genes determined in this work can be used for dose optimization in adjuvant tumor therapy using Silibinin. Further investigations with a broader spectrum of NSCLC cells and other cancer cell lines are required to generalize preliminary findings of this study.

**CONCLUSION**

Silibinin inhibits NSCLC cell viability and motility in a dose-dependent manner. Thereby, SIL IC50 of different NSCLC cells does not correlate with their transcriptomic EMT signature. Instead, it was found that SIL dose–response correlates with the expression level of a network of 144 cell cycles, stress responsive and survival proteins including some well known targets of STAT3. It was shown that selected downstream targets of STAT3 including BRCA1, BIRC5, FOXM1 have impact on Silibinin dose response.
ACKNOWLEDGEMENTS

The authors are grateful to Dr. Marco Binder and Jürgen Beneke for helping with optimization and implementation of transfection assays, and to Katharina Jechow and Ellen Wiedtke for their support in cell culturing and microscopic imaging.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding
Jagan Mohan Kaipa was supported by the DKFZ-MOST Grant c-160. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures
The following grant information was disclosed by the authors: DKFZ-MOST: c-160.
Competing Interests
The authors declare there are no competing interests.

Author Contributions
• Jagan Mohan Kaipa and Evgeny Gladilin conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
• Vytaute Starkuviene conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
• Holger Erfle and Roland Eils conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability
The following information was supplied regarding data availability:
The raw data are available in Tables S1–S4.

Supplemental Information
Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.10373#supplemental-information.

REFERENCES


