The Role and Molecular Mechanism of SLC34A2 in Stemness Maintenance of $CD44^{+}CD166^{+}$ Lung Cancer Stem Cells (LCSCs) with AT-II cells' characteristics

Yan Yang
Sichuan University State Key Laboratory of Biotherapy
Qiang Pu
Sichuan University West China Hospital
Hu Liao
Sichuan University West China Hospital
Yue Yuan
Sichuan University State Key Laboratory of Biotherapy
Xueting Hu
Sichuan University State Key Laboratory of Biotherapy
Xinmei Luo
Sichuan University State Key Laboratory of Biotherapy
Qianqian Jiang
Sichuan University State Key Laboratory of Biotherapy
Xiaolan Su
Sichuan University State Key Laboratory of Biotherapy
Wen Zhu (zhuwen@scu.edu.cn)
Sichuan University State Key Laboratory of Biotherapy https://orcid.org/0000-0002-9827-7766

Research Article

Keywords: CD44+CD166+ lung cancer stem cells, SLC34A2, Stemness maintenance, Type II alveolar epithelium cells, PI3K/AKT/STAT3/Sox2 axis, Initiation of cancer

Posted Date: March 31st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-368100/v1

Abstract

Background

Evidence showed some non-small cell lung cancers (NSCLCs) had the type II alveolar epithelium cells'(AT-II cells) characteristics, and AT-II cells were a kind of original stem cells of NSCLCs. But how AT-II cells malignantly transformed into NSCLCs was unclear. Recent evidence indicated SLC34A2 was critical in the development of AT-II cells, and SLC34A2 might be a new gene in the initiation of NSCLCs. However, whether SLC34A2 participated in the malignant transformation of AT-II cells remained unknown. The exact role and mechanism of SLC34A2 in the initiation of NSCLCs needed to be further investigated.

Methods

The expression of Napi-IIb (encoded by SLC34A2) in the NSCLC cells was compared with that in AT-II cells using immunohistochemistry (IHC). Also coexpression of CD44 and CD166 was detected in these NSCLCs tissues by IHC. Then the $CD44^{+}CD166^{+}$ cells were sorted from lung tumor spheres by FACS. They were assessed by sphere, proliferation and tumorgenicity assay. Besides, their expression of surfactants C(SP-C) was stained by IHC. Next, the role and mechanism of SLC34A2 in $CD44^{+}CD166^{+}$ lung cancer stem cells were explored by siRNA-mediated SLC34A2 knockdown, related pathway pharmacological inhibition or activation. In vitro findings were furtherly validated *in vivo* and NSCLCs samples.

Results

The expression of SLC34A2 was downregulated in NSCLCs cells compared with AT-II cells in clinic samples. Then the $CD44^{+}CD166^{+}$ population was identified as $CD44^{+}CD166^{+}$ lung cancer stem cells (LCSCs). And LCSCs showed abundant expression of SP-C, the hallmark of AT-II cells. Higher expression of SLC34A2 was found in LCSCs compared to their origin NSCLC cells. Additionally, the expression of SLC34A2 was decreased after LCSCs were differentiated, and the morphology of the differentiated cells from LCSCs was similar to their origin NSCLC cells. Knockdown SLC34A2 made declined abilities of self-renewal, drug-resistance, migration and invasion *in vitro* as well as tumorigenicity *in vivo* in LCSCs. And SLC34A2 could maintain stemness of LCSCs via PI3K/AKT/STAT3/Sox2 axis. Besides, the connection between SLC34A2 maintaining stemness of lung cancer stem cells and PI3K/AKT/STAT3/Sox2 axis was also validated *in vivo* and in clinic samples.

Conclusions

For the first time, we illustrated the expression of SLC34A2 was downregulated in NSCLCs cells compared with AT-II cells. We discovered the downregulated expression of SLC34A2 performed a vital role in the malignant transformation of AT-II cells into NSCLCs. And SLC34A2 could maintain stemness of

$CD44^+CD166^+$ lung cancer stem cells, which were with AT-II cell's characteristic, via PI3K/AKT/STAT3/Sox2 axis. It had important significance in the revelation of a new mechanism for the initiation of NSCLCs.

Introduction

Non-small cell lung cancers (NSCLCs) accounted for 80% -85% of lung cancers [1]. As their non-ideal diagnosis at the early stage and treatment, the 5-year survival of NSCLCs was still less than 20% [2]. So, refined investigation on the new mechanism under the origin of NSCLCs will strongly promote its early diagnosis and provide effective methods for its targeted therapy.

The cancer stem cell hypothesis stated that the origin of cancer was fuelled by cancer stem cells (CSCs), and the deregulated self-renewal pathways in CSCs contributed to the initiation of cancers [3, 4]. The genetic alterations might make stem cells to malignantly transform into cancer stem cells, then cancer stem cells were differentiated into cancer cells [5–7]. Accumulated studies showed that lung cancer stem cells promoted the origin of NSCLCs [8, 9]. And some genes and pathways involved in embryonic development were found to participate in retaining stemness of lung cancer stem cells [10]. In recent years, as the complex anatomical features of the lung, many types of stem cells were found in different anatomical sites of the lung [10]. Diverse lung stem cells or progenitor cell populations with the abilities of self-renewal and differentiation, were also found to have the ability to drive lung oncogenesis in different contexts [10]. For example, in the proximal lung, the tracheal basal cells had been proposed to be the cells of origin for NSCLCs; in the bronchoalveolar junction, the bronchioalveolar stem cells (BASCs) might be able to give rise to NSCLCs; at the end of the airway, the type II alveolar epithelial cells(AT-II cells) appeared to progress into NSCLCs [9]. However, how these stem cells malignantly transformed were rarely understood until now.

The alveolar was very important to maintain the physiological function of the lung. The alveolar was the main place for gas exchange, while it was highly susceptible to injury as it was directly exposed to the external environment [1]. And the alveolar was a major site of NSCLCs [11, 12]. The alveolar consisted of the type I and type II alveolar epithelial cells (AT- I and AT-II) [13]. And the AT-II cells, which occupied about 4% of the alveolar surface, were cubic and located in the periphery of the alveoli [13]. The AT-II cells were critical to the physiological function of the alveolar, and they could prevent alveolar collapse for gas exchange by secreting surfactants [12]. Among the surfactants, the surfactant protein C(SP-C) was restricted exclusively to the AT-II cells and was identified as the hallmark of AT-II cells [12, 14]. More noteworthy, the AT-II cells could be intermittently activated to self-renewal and differentiate into the AT-I cells for restoring alveolar in physiological condition and during lung injury, so the AT-II cells were identified as stem cells in alveolar [11, 12]. Recently, some studies focused on the relationship between the malignant transformation of the AT-II cells and the origin of NSCLCs. The early histopathological studies uncovered some cells in NSCLCs tissues exhibited the features of AT-II cells, such as the presence of SP-C in the cell membrane [15, 16]. Furthermore, in adult mouse lung, AT-II cells appeared to progress into NSCLCs by Kras mutation or by Kras mutation combined with Notch inhibition [17–19]. These results

indicated that AT-II cells, the stem cells in alveolar, were a kind of original stem cells of NSCLCs. But how the AT-II cells malignantly transformed into NSCLCs was not clear until now.

SLC34A2, encoding a type 2b sodium-dependent phosphate transporter(Napi-IIb), was responsible for transporting extracellular phosphorus ions into cells[20, 21]. It was critical not only to the normal function of the lung but also to lung development. Among various organs, the highest expression of SLC34A2 was found in the lung and fetal lung[22]. And during the development of the human fetal lung, the expression of SLC34A2 was up-regulated[22]. More importantly, the latest study revealed that in terminal lung epithelium, SLC34A2 was specifically expressed in mouse mature AT-II cells by single-cell RNA sequencing, and the expression of SLC34A2 was gradually upregulated during alveolar epithelial precursor cells differentiating into the AT-II cells[14]. Similarly, Jacob A et.al also found the expression of SLC34A2 was gradually increased, when pluripotent stem cells (iPS) differentiated into the mature AT-II cells[23]. These results indicated the key role of SLC34A2 in the development of the AT-II cells. Tumorigenesis resembled the reverse process of embryogenesis[3]. Kopantzev EP et.al found the expression of SLC34A2 in 12/12 NSCLCs tissues was significantly down-regulated[24]. Our previous studies further confirmed that expression of SLC34A2 was significantly downregulated in 59/90 NSCLCs tissues and 6 NSCLCs cell lines, and subsequently, the abilities of proliferation, invasion and migration were significantly inhibited in NSCLCs cells after SLC34A2 was overexpressed[25, 26]. These results showed that the downregulated expression of SLC34A2 was involved in the initiation of NSCLCs. But the role and mechanism of SLC34A2 in the initiation of NSCLCs needed to be further investigated. Cancer stem cell hypothesis thought that abnormal expression of genes or pathways involved in embryonic development and self-renewal might lead to malignant transformation of stem cells[7]. Therefore, we further guessed the downregulated expression of SLC34A2 might be involved in the malignant transformation of AT-II cells into NSCLCs. However, until now little is known about it.

Therefore, in our present study, we aimed to investigate the relationship between downregulated expression of SLC34A2 and the malignant transformation of AT-II cells. And the role and mechanism of SLC34A2 in the initiation of NSCLCs would be further investigated. Firstly, we compared the expression of SLC34A2 in NSCLCs cells and AT-II cells in clinic samples. Then we identified $CD44^+CD166^+$ lung cancer stem cells (LCSCs) from NSCLCs cell lines and tissue, and found LCSCs were with AT-II cell's characteristic. Furtherly we explored the role and molecular mechanism of SLC34A2 in stemness maintenance of $CD44^+CD166^+$ lung cancer stem cells (LCSCs). For the first time, our studies demonstrated SLC34A2 played an important role in the malignant transformation of AT-II cells into NSCLCs. And our results shed light on the revelation of the new mechanism for the initiation of NSCLCs.

Materials And Methods

NSCLCs Sample Collection

A total of 57 NSCLC tissues and their adjacent tissues were obtained from the Department of Thoracic Surgery, West China Hospital, Sichuan University. This study was performed with the approval of the

Medical Ethical Committee of West China Hospital, Sichuan University. All patients were diagnosed with primary NSCLC and did not show other tumor occurrences. The detailed information of clinical parameters was shown in Additional files 1.

Quantitative Real-time RT-PCR Analysis

Differently treated cells were respectively washed with cold PBS and then extracted total RNA using Trizol (Invitrogen, USA) according to the manufacturer's protocol. Extracted total RNA was reverse transcribed into cDNA using the PrimrScript RT reagent Kit with gDNA Eraser (TaKaRa, Japan). As for the protocol of the quantitative real-time PCR analysis, it **w**as described in Additional files 1.

Immunohistochemistry (IHC)

Immunohistochemistry was performed in NSCLC tissues, their adjacent tissues and mouse xenografts. All of the specimens were embedded in paraffin and cut as the sections of 4 μm thick, and then carried out the standard histological analysis. As for the protocol of immunohistochemistry, it was described in Additional files 1.

Animal and Cell Culture

A detailed description was made in Additional files 1.

Sphere Formation and Propagation

Single-cell suspensions were obtained from NSCLC tissue and NSCLC cell lines respectively. After mechanical and enzymatical disaggregation, cells were plated in serum-free medium, and the details were Additional files 1.

Colony Formation Assay

A detailed description was made in Additional files 1.

Xenotransplantation

A detailed description was made in Additional files 1.

Flow Cytometry Analysis and Cytofluorometric Cell Separation

A detailed description was made Additional files 1.

Immunofluorescence

A detailed description was made Additional files 1.

Hematoxylin-Eosin (H&E) Staining

Hematoxylin-eosin (H&E) staining was performed in mouse xenografts. All the specimens were embedded in paraffin and cut as sections of 4 μm thick, then they were stained with hematoxylin-eosin

(H&E) according to standard histopathological techniques.

Differentiation of Stem Cell Progeny

A detailed description was made in Additional files 1.

Cell growth assay

A detailed description was made in Additional files 1.

Cell Viability Assay

A detailed description was made in Additional files 1.

Lentiviral Transduction

A549-LCSCs and H460-LCSCs were dissociated and spin-infected with 0.5 ml of NC-si-RNA (Shanghai Hanbio Biotechnology Co. Ltd, China) or *SLC34A2*-si-RNA lentiviral knockdown package (Shanghai Hanbio Biotechnology Co. Ltd, China). For both systems, cells were infected with lentiviral media at the MOI value of 40 overnight in a 37°C incubator. The transduction efficiency was evaluated by GFP expression. Stable clones with the GFP expression was > 99%, were then selected using puromycin. Note: LCSCs was short for $CD44^{+}CD166^{+}$ lung cancer stem cells.

Western Blot

A detailed description was made in Additional files 1.

Transwell invasion /migration assay

A detailed description was made in Additional files 1.

Chromatin immunoprecipitation analysis

A detailed description was made in Additional files 1.

The Cancer Genome Atlas (TCGA) analysis

A detailed description was made in Additional files 1.

Statistical analysis

All the data was analyzed using SPSS 13.0 statistical software (SPSS Inc, USA). The graphed data presented as mean ± SEM used two-tailed Student's t-tests. $p < 0.05$ was designated as statistically significant. Significances are ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

Results

Down-regulated Expression of SLC34A2 Was Involved in Malignant Transformation of AT-II cells in Clinic Samples

Although it had been reported that the expression of SLC34A2 in NSCLCs was significantly lower than that in their adjacent tissues[24–26], yet there had no direct report about the relative expression of SLC34A2 between NSCLCs and AT-II cells. Therefore, we first compared the expression of SLC34A2 in NSCLCs and AT-II cells in clinic samples. 57 poorly or moderately differentiated NSCLCs tissue samples were collected. And the downregulated expression of SLC34A2 was found in 45/57 collected poorly or moderately differentiated NSCLCs samples compared with their adjacent tissues (Fig. 1A). Namely, in most NSCLCs samples, the expression of SLC34A2 was found to be significantly downregulated compared with their adjacent tissues, which is consistent with our previous reports [25, 26]. Then 45 NSCLCs samples with the down-regulated expression of SLC34A2 were chosen for further study. Classically, SP-C was identified as the hallmark of AT-II cells [12, 14]. We detected the expression of SP-C by immunohistochemistry in the 45 NSCLCs tissues with down-regulated expression of SLC34A2. Cells with positive expression of SP-C were present in 18/45 NSCLCs samples (Additional File 2), indicating these 18 NSCLCs samples might be originated from the malignant transformation of AT-II cells. Secondly, the 18 NSCLCs samples with positive expression of SP-C and their adjacent tissues were chosen for further study. Referring to the previous literature [33], the corresponding adjacent tissues of the 18 NSCLCs were stained with both Napi-IIb and SP-C by double staining technology in serial sections to detect the expression of Napi-IIb in AT-II cells. The AT-II cells were cubic and located in the periphery of the alveoli [13]. In the sections, the cells with positive expression of SP-C were mostly found in the periphery of the alveoli and cuboidal. Similarly, the cells with positive expression of Napi-IIb were also mostly found in the periphery of the alveoli and their shape was cuboidal (Fig. 1B). Not only SP-C, but also SLC34A2 was identified as the hallmark of the AT-II cells recently [14]. Thus, we concluded the cells with positive expression of both Napi-IIb and SP-C were AT-II cells. Then we found that the staining intensity of Napi-IIb in the NSCLCs cells on the tumor tissue section was weaker than that in the AT-II cells on the adjacent tissue section (Fig. 1C). All these results showed the expression of SLC34A2 was downregulated in the NSCLCs cells compared with the AT-II cells. And it indicated down-regulated expression of SLC34A2 was involved in the malignant transformation of AT-II cells.

The Expressions of CD44 and CD166 Were Evaluated in NSCLCs Samples and Spheres from NSCLCs Cell Lines and NSCLCs Tissue

The primary AT-II cells were difficult to culture and maintain in a well-differentiated state [12], so it was difficult to explore how AT-II cells malignantly transformed into NSCLCs directly. According to the cancer stem cell hypothesis stated that the genetic alterations might make stem cells to malignantly transform into cancer stem cells, then cancer stem cells were differentiated into cancer cells [5]. Therefore, we supposed AT-II cells might malignantly transform into lung cancer stem cells and then differentiate into NSCLCs cells. And we intended to explore the role of SLC34A2 in the malignant transformation of AT-II cells through the lung cancer stem cells which were with AT-II cells' characteristic.

The method for sorting lung cancer stem cells by single surface marker, such as CD133, CD326, CD44 and CD166 was controversial [28]. Researchers pointed out that it would be better to combine more than one marker to isolate CSCs from NSCLCs [29]. Evidence showed the expression of CD44 or CD166 was

related to stem cell characteristics of AT-II cells [30, 31]. Thus, we speculated the lung cancer stem cells with the surface markers of CD44 and CD166 might be related to the AT-II cells. In an attempt to preliminarily identify the markers of lung cancer stem cells, 18 primary NSCLC tissue samples with the down-regulated expression of SLC34A2 compared with their adjacent tissues and positive expression of SP-C were chosen to study. They were double-stained with the surface marker CD44 and CD166 by IHC in the serial section. Some cells were detected to show both positive expressions of CD44 and CD166 in 18 out of 18 NSCLC samples (include 12 lung adenocarcinoma and 6 lung squamous cell carcinoma) (Additional File 3, Fig. 2A). As a result, it indicated that the expression of $CD44^+CD166^+$ was related to AT-II cells in NSCLCs specimens.

Furthermore, 3 NSCLC cell lines (A549, H460, and H1299) and 1 NSCLC tissue were cultured in the serum-free medium to enrich CSCs. After 2 weeks, some cells in 3 NSCLC cell lines and 1 NSCLC tissue could form spheres (Fig. 2B). And these spheres were found to possess significantly higher colony formation efficiency compared with their corresponding NSCLC cell lines and differentiated cells respectively (Fig. 2C and 2D). Then the injection of as low as 10^5 sphere cells (except H1299 spheres) was uncovered to consistently produce tumor xenografts in 3–5 weeks (Fig. 2E and Additional File 4). However, 10^6 NSCLCs cells (A549 and H460) and the differentiated cells from NSCLC tissue spheres could not generate tumor xenografts in 5 weeks (Additional File 4). All demonstrated that the spheres from NSCLC cell lines (A549 and H460) and NSCLC tissue had a higher capability of proliferation *in vitro* and tumorigenicity *in vivo*. And it indicated that these spheres contained lung cancer stem cells. Given that H1299 spheres failed to generate tumor xenografts in about 5 weeks, which pointed out the weaker CSCs characteristics of H1299, so we only took the spheres from A549, H460 and NSCLC tissue for the following study. After that, the higher proportion of $CD44^+CD166^+$ cells was found in spheres than in their corresponding NSCLC cell lines (A549 and H460) or differentiated cells by FACS (Fig. 2F). It furtherly showed that $CD44^+CD166^+$ could be the roust markers of lung cancer stem cells in spheres from NSCLCs cell lines and NSCLCs tissue.

$CD44^+CD166^+$ Spheres from NSCLC Tissue and NSCLC Cell Lines Were Identified as Lung Cancer Stem Cells with AT-II cells' Characteristic

To identify whether $CD44^+CD166^+$ could be a marker for lung cancer stem cells with AT-II cells' characteristic, firstly the fractions of $CD44^+CD166^+$, $CD44^+CD166^-$, $CD144^-CD166^+$, and $CD44^-CD166^-$ were respectively sorted from the spheres by FACS. Subsequently, the strongest ability of sphere formation *in vitro* and tumor formation *in vivo* was detected in $CD44^+CD166^+$ cells than $CD44^+CD166^-$ and $CD144^-CD166^+$ cells (Fig. 3A,3B and Additional File 5). Moreover, the highest potential of proliferation was found in transplants of $CD44^+CD166^+$ cells by Ki-67 assay (Fig. 3C). And the mice xenografts derived by $CD44^+CD166^+$ spheres were found to show positive expression of CD44 and CD166 (Fig. 3F). Besides, $CD44^+CD166^+$ spheres had upregulated expression of the genes involved in stemness, including *β-catenin*, *Sox2, Bmi-1, Oct-4, Nanog* and *CXCR4*, compared with their corresponding NSCLC cell lines (A549 and H460) (Fig. 3E). And $CD44^+CD166^+$ spheres also had increased expression of

the genes related to drug resistance, such as *ABCG1* and *Notch3*, as well as the enhanced viability against cisplatin and the higher value of IC_{50} (Fig. 3E and Additional File 6).

After being cultured in serum-medium, $CD44^+CD166^+$ cells could differentiate into the cells which had similar morphological features with their original NSCLCs cells (Fig. 3D). At the same time, the expression of the genes involved in stemness, including *β-catenin*, *Sox2, Bmi-1, Nanog* and *CXCR4*, was notably reduced after $CD44^+CD166^+$ spheres were differentiated (Fig. 3E). Similarly, the expression levels of the genes related to drug resistance, e.g. *ABCG1* and *Notch3*, as well as the viability against cisplatin and value of $IC_{50,}$ were also notably declined in differentiated $CD44^+CD166^+$ spheres (Fig. 3E and Additional File 6). Thus, we demonstrated a poorly differentiated state of $CD44^+CD166^+$ spheres.

The common method used for identifying the AT-II cells was to detect the expression of the surfactant protein C(SP-C) at present [14]. To explore whether $CD44^+CD166^+$ spheres were originated from AT-II cells, we detect the expression of SP-C in $CD44^+CD166^+$ spheres. It was consistent with our expectations that $CD44^+CD166^+$ spheres showed abundant expression of SP-C (Fig. 3G,3H). Also the mice xenografts derived by $CD44^+CD166^+$ spheres showed abundant expression of SP-C (Fig. 3I). These showed that $CD44^+CD166^+$ spheres were with AT-II cells' characteristic, and indicated that $CD44^+CD166^+$ spheres might be originated from malignantly transformed AT-II cells.

Collectively, $CD44^+CD166^+$ spheres were identified as $CD44^+CD166^+$ lung cancer stem cells with the AT-II cells' characteristic. In the following study, we named $CD44^+CD166^+$ lung cancer stem cells as LCSCs (Lung Cancer Stem Cells) for short.

SLC34A2 **was Discovered to Maintain Stemness of $CD44^+CD166^+$ Lung Cancer Stem Cells** in vitro **and** in vivo

Next, we tested the role of SLC34A2 played in LCSCs. The higher mRNA and protein expression levels of SLC34A2 were found in LCSCs than that in their corresponding NSCLC cell lines (Fig. 4A). Intriguingly, the differentiated cells that were derived from LCSCs had the lower expression of SLC34A2, compared with their LCSCs counterparts (Fig. 4A). The results above suggested that SLC34A2 might play a crucial role in maintaining stemness in LCSCs.

To verify this hypothesis, a lentiviral-based approach was used to knockdown *SLC34A2* in LCSCs with si-RNAs, and the knockdown efficiency was confirmed by qRT-PCR and western blotting (Fig. 4B). Following, the weaker abilities of sphere formation, colony formation *in vitro* were detected in Si-*SLC34A2*-LCSCs than Si-NC-LCSCs and LCSCs (Fig. 4C and 4D). Importantly, when implanted into the flank of nude mice, the tumorigenicity of si-*SLC34A2*-LCSCs was significantly lower than NC-LCSCs and LCSCs in vivo (Figs. 4E and Additional File 7). Moreover, compared with tumor xenografts of si-NC-LCSCs and LCSCs, tumor xenografts of si-SLC34A2-LCSCs also had the weaker ability of proliferation by Ki-67 staining (Figs. 4F). Then we tested the expression of genes involved in stemness, and found the mRNA expression level of the genes related to stemness - Sox2 was declined after SLC34A2 knockdown in A549-LCSCs and

H460-LCSCs (Fig. 4G). Also the mRNA expression level of Oct-4, Bmi-1, CXCR4, Nanog and β-catenin was decreased after SLC34A2 knockdown in A549-LCSCs (Fig. 4G).

Collectively, our results demonstrated that SLC34A2 was necessary to maintain the stemness of LCSCs.

SLC34A2 was Discovered to Maintain Stemness of $CD44^+CD166^+$ Lung Cancer Stem Cells by PI3K/AKT/STAT3/Sox2 Axis

Following, we intended to search the molecular mechanism under SLC34A2 maintaining stemness of LCSCs. As the mRNA expression level of *Sox2* was reduced after SLC34A2 knockdown both in A549-LCSCs and H460-LCSCs, therefore, we suggested that *Sox2* might play an important role in the process that *SLC34A2* maintaining stemness of LCSCs. Then the declined protein expression level of *Sox2* was confirmed in si-*SLC34A2*-LCSCs than Si-NC-LCSCs and LCSCs (Fig. 4H). Moreover, the NaPi-IIb protein, encoded by *SLC34A2*, was located in the cell membrane [20], thus SLC34A2 might regulate the transcription of Sox2 in nuclear by its downstream factors. Later, the reduced expression of key proteins in PI3K/AKT (PI3K and p-AKT) and STAT3 pathway(p-STAT3) was discovered in Si-*SLC34A2*-LCSCs than Si-NC-LCSCs and LCSCs (Fig. 5A). Additionally, significantly lower expression of key proteins (p-NF-KB and p-STAT3) in STAT3 pathway was also found in LCSCs after inhibition of PI3K/AKT pathway (LY294002) or AKT(MK2206), while inhibition of STAT3 pathway (WP1066) did not affect the expression of the key proteins (PI3K and p-AKT) in PI3K/AKT pathway in LCSCs (Fig. 5B-5E and Additional File 8). These results showed that the PI3K/AKT/STAT3 axis was located downstream of *SLC34A2* in LCSCs.

After that, we aimed to find whether *SLC34A2* regulated the transcription of *Sox2* by the PI3K/AKT/STAT3 axis. Similar to the effect of SLC34A2 knockdown, reduced mRNA and protein expression levels of *Sox2* were detected in LCSCs after inhibition of the PI3K/AKT/STAT3 axis (Fig. 5F). While significantly increased mRNA and protein expression levels of *Sox2* were found in Si-*SLC34A2*-LCSCs after activation of the PI3K/AKT/STAT3 axis (Fig. 5G). The above showed that *SLC34A2* could regulate the transcription level of *Sox2* by the PI3K/AKT/STAT3 axis. Furthermore, binding sites of STAT3 to the promoter region of Sox2 were predicted via the bioinformatics prediction website (Genecards, JASPAR and Patch), and one site was confirmed by CHIP-PCR according to the reference [32] (Fig. 5H and 5I). Then, the CHIP-qPCR result uncovered the binding probability of STAT3 to the promoter region of Sox2 in LCSCs was decreased after SLC34A2 was interfered (Fig. 5J). Collectively, the above showed that SLC34A2 prompted STAT3 to bind to the promoter region of Sox2 via PI3K/AKT pathway in LCSCs.

Finally, we confirmed whether SLC34A2 maintained stemness of LCSCs via the PI3K/AKT/STAT3/Sox2 axis. Similar to the effect of SLC34A2 knockdown, the reduced stemness was found in LCSCs after inhibition of the PI3K/AKT/STAT3/Sox2 axis, such as the declined abilities of sphere formation, colony formation, growth, drug-resistance, migration, and invasion, as well as improved expression of differentiation markers CK8 and CK18 (Fig. 6A-6G and Additional File 9). While these declined abilities in si-*SLC34A2*-LCSCs, such as sphere formation, colony formation, growth, migration and invasion, as well as declined expression of differentiation markers CK8 and CK18, could be rescued by activation of the

PI3K/AKT/STAT3/Sox2 axis (Fig. 6A-C, E-G). The ability of drug-resistance was also discovered to be largely strengthened in A549-si-*SLC34A2*-LCSCs after the PI3K/AKT/STAT3/Sox2 axis was activated (Fig. 6D). These results showed that SLC34A2 maintained stemness of LCSCs via the PI3K/AKT/STAT3/Sox2 axis *in vitro.*

The Connection between SLC34A2 Maintaining Stemness of Lung Cancer Stem Cells and PI3K/AKT/STAT3/Sox2 axis Was Validated in vivo and in clinic samples

At last, we validated the connection between SLC34A2 maintaining stemness of lung cancer stem cells and PI3K/AKT/STAT3/Sox2 axis in vivo and in clinic samples. In vivo, the weaker positive expression of key proteins in the PI3K/AKT/STAT3/Sox2 axis (PI3K, p-AKT, p-STAT3 and Sox2) was found in xenografts derived from si-SLC34A2-LCSCs than si-NC-LCSCs (Fig. 6H). Combined with our *in vivo* finding that Si-*SLC34A2*-LCSCs had reduced ability of tumorigenicity than Si-NC-LCSCs, these results indicated the connection between *SLC34A2* maintaining stemness of lung cancer stem cells and the activation of the PI3K/AKT/STAT3/Sox2 axis *in viv*o.

As we mentioned in the first part, we found the downregulated expression of SLC34A2 in 45/57(most) poorly or moderately differentiated NSCLCs samples compared with their adjacent tissues (Fig. 1A). What's interesting was that we found the expression of SLC34A2 in LCSCs, which were enriched and originated from NSCLCs cells, was significantly higher than that in NSCLC cells in vitro (Fig. 4A). These results suggested that although the expression of SLC34A2 was significantly downregulated in NSCLC cells compared with normal cells, the lung cancer stem cells (a subset of NSCLCs) had relatively higher expression of SLC34A2 compared with other NSCLCs cells. To furtherly detect the connection between SLC34A2 maintaining stemness of lung cancer stem cells and PI3K/AKT/STAT3/Sox2 axis in clinic samples, the 45 NSCLCs samples, which had the down-regulated expression of SLC34A2 compared with their adjacent tissues, were chosen for further study. And later they were divided into two groups according to their IHC score of Napi-IIb: the group with relatively higher expression of SLC34A2(the $SLC34A2^{High}$ group) (IHC score of Napi-IIb $\geq$ 6) and the group with relatively lower expression of SLC34A2(the $SLC34A2^{Low}$ group) (IHC score of Napi-IIb < 6). Subsequently, data analysis displayed that the proportion of NSCLCs samples with high T stage had no significant difference between the two groups (Fig. 6I). It indicated that the downregulated expression of SLC34A2 might be associated with the early origin of NSCLCs. Besides, compared with the $SLC34A2^{Low}$ group, the $SLC34A2^{High}$ group had a higher proportion in NSCLCs samples with lymphatic metastasis and high expression of p-AKT (Fig. 6J and 6K). However, the proportion of NSCLCs samples with high expression of PI3K, p-STAT3 and Sox2 had no difference between the two groups (Fig. 6J and 6K). Cancer stem cells had a stronger ability of invasion and metastasis than cancer cells [4]. Therefore, these results preliminary suggested the connection between SLC34A2 maintaining stemness of lung cancer stem cells and the activation of AKT in the collected NSCLCs samples.

Besides, the connection between SLC34A2 maintaining stemness of lung cancer stem cells and the activation of AKT and STAT3 was also preliminarily found in the datasets of the TCGA database

(Additional File 10).

Discussion

The cancer stem cell hypothesis assumed cancer stem cells were originated from stem cells [4]. Some studies on histopathological sections and mouse models indicated that AT-II cells, the stem cells in alveolar, were the origin cells of NSCLCs [15–19]. But how AT-II cells malignantly transformed into NSCLCs is still unclear now. Extensive researches suggested that abnormal expression of genes or pathways involved in embryonic development and self-renewal might cause stem cells to malignantly transform into cancer stem cells and promote the initiation of cancers [3, 6, 7]. Recent research found SLC34A2 was specifically expressed in mouse mature AT-II cells and the expression of *SLC34A2* was significantly upregulated during alveolar precursor cells differentiating into the AT-II cells [14]. It showed upregulated expression of SLC34A2 played a key role in the development of AT-II cells. Tumorigenesis resembled the reverse process of embryogenesis [3]. Yet, our previous studies and Kopantzev EP et.al both found the expression of SLC34A2 in most NSCLCs tissues were significantly down-regulated compared with their adjacent tissues, while the abilities of proliferation, invasion and migration were significantly inhibited in NSCLCs cells after SLC34A2 was overexpressed [24–26]. These results showed downregulated expression of SLC34A2 was involved in the initiation of NSCLCs. Therefore, according to the cancer stem cell hypothesis, we furtherly supposed the downregulated expression of SLC34A2 might be related to the malignant transformation of AT-II cells.

Although it had been reported that the expression of SLC34A2 in NSCLCs was significantly lower than that in their adjacent tissues [24–26], there was no direct evidence about the relative expression of SLC34A2 in NSCLCs cells and AT-II cells. Therefore, for the first time, we preliminary investigated the relationship between downregulated expression of SLC34A2 and malignant transformation of AT-II cells in clinic samples. The significantly down-regulated expression of SLC34A2 was found in most (45/57) NSCLCs samples from collected NSCLCs tissues compared with their adjacent tissues. But it was hard to compare the relative expression of SLC34A2 in NSCLCs cells and AT-II cells in clinic samples. Mitsuyoshi Hashimoto, et.al determined the expression of Napi-IIb (encoded by SLC34A2) in rat AT-II cells by in situ hybridization. He used serial sections of rat lung tissues to stain Napi-IIb and SP-C. The purpose of staining SP-C was to identify the AT-II cells in rat lung tissue. If the cells showed positive expression of SP-C, and were in the periphery of the alveoli and cubic, they were considered as the AT-II cells. Interestingly, Mitsuyoshi Hashimoto, et.al found the cells which showed positive expression of Napi-IIb, also located in the periphery of the alveoli and cubic. And cells with positive expression of Napi-IIb had similar position and morphology to the cells with positive expression of SP-C. So Mitsuyoshi Hashimoto, et.al also identified the cells with positive expression of Napi-IIb as AT-II cells in rat lung tissue [27]. Similarly, in our study, we detected the expression of Napi-IIb and SP-C in AT-II cells referring to the previous methods [27]. The serial sections of 18 adjacent lung cancer tissues were stained Napi-IIb and SP-C by immunohistochemical staining. Some cells showed positive expression of SP-C in adjacent NSCLCs tissue sections and located mostly in the periphery of the alveoli, cubic. And the cells with positive expression of Napi-IIb had the similar position and morphology to the cells with positive

expression of SP-C. According to the literature [27], we considered the cells with positive expression of Napi-IIb in the periphery of the alveoli and cubic as the AT-II cells. Recently, SLC34A2 was identified as the hallmark of the AT-II cells [14]. So, it furtherly confirmed our findings that in the periphery of the alveoli, the cells with positive expression of Napi-IIb were indeed the AT-II cells. Then the weaker staining intensity of Napi-IIb was found in the NSCLCs cells in the 18 tumor tissue sections compared with the AT-II cells in their adjacent tissue sections in our present study. Namely, the expression of SLC34A2 in the NSCLCs cells was decreased compared with the AT-II cells. For the first time, our results indicated that the downregulated expression of SLC34A2 might lead AT-II cells to malignantly transform into NSCLCs from the perspective of histology.

As it was difficult to culture and maintain primary AT-II cells in a well-differentiated state, as well as faithfully recapitulate the morphologic, functional, molecular, proteomic, and/or lipidomic markers of AT-II phenotype in stable cell lines [12], the research on how AT-II cells were malignantly transformed into NSCLCs was greatly hindered before. According to the cancer stem cell hypothesis, the genetic alterations might make stem cells to malignantly transform into cancer stem cells, then cancer stem cells were differentiated into cancer cells [5]. We supposed AT-II cells might be malignantly transformed into lung cancer stem cells and then differentiated into NSCLCs cells. Therefore, we intended to explore the role and mechanism of SLC34A2 in the malignant transformation of AT-II cells by the lung cancer stem cells with AT-II cells' characteristic. Evidence showed the expression of CD44 or CD166 was related to stem cell characteristics of AT-II cells [30, 31]. Thus, we speculated the lung cancer stem cells with the surface markers of CD44 and CD166 might be related to AT-II cells. Then we preliminarily explored the connection between the expression of CD44 and CD166 and the AT-II cells in 18 NSCLCs samples. Histological analysis revealed that in the 18 NSCLCs tissues with downregulated expression of SLC34A2 and positive expression of SP-C, some cells showed coexpression of CD44 and CD166. These results suggested that the expression of $CD44^{+}CD166^{+}$ was related to AT-II cells in NSCLCs specimens. However, neither CD44 nor CD166 was associated with TNM stage and differentiation in NSCLCs samples, which might most likely due to the limited number of NSCLC tissues. And no prognosis data was available for these NSCLCs samples, so we were unclear whether the expression of $CD44^{+}CD166^{+}$ correlated with the prognosis of NSCLCs. In the future, we will add the number of NSCLCs samples and collected their prognosis data to refine this part.

Futherly, we sorted the $CD44^{+}CD166^{+}$ lung cancer stem cells from NSCLCs tissue and cell lines. Spheres, which possessed the higher abilities of colony formation and tumorigenic potential *in vivo*, were enriched from NSCLCs tissue and cell lines by cultivating in a serum-free medium. Then the higher proportion of $CD44^{+}CD166^{+}$ was tested in spheres than that in their corresponding NSCLCs cell lines and differentiated cells. Subsequently, $CD44^{+}CD166^{+}$, $CD44^{+}CD166^{-}$, $CD44^{-}CD166^{+}$ and $CD44^{-}CD166^{-}$ cells were sorted from spheres by FACS. And the $CD44^{+}CD166^{+}$ cells were found to have the strongest abilities of sphere formation and tumorigenesis than other fractions respectively. Besides, $CD44^{+}CD166^{+}$ cells had a poor differentiation state and enhanced ability of drug resistance against cisplatin, as well as the increased expression of genes involved in stemness, compared with their adherent counterparts *in vitro*. These

results identified the $CD44^+CD166^+$ population as lung cancer stem cells. Next, we stained surfactant protein C(SP-C) in $CD44^+CD166^+$ spheres to determine whether $CD44^+CD166^+$ spheres had AT-II cells' characteristic. SP-C was one kind of surfactant proteins secreted by the AT-II cells [12]. SP-C was not only essential for the physical structure of the surfactant, but also the formation and stability of the surface-active monolayer at the air-water interface [13]. In the adult, only the expression of SP-C was restricted exclusively to the AT-II cells, and SP-C was identified as the hallmark of the AT-II cells [12, 14]. Therefore, the common method used for identifying the AT-II cells was to detect the expression of SP-C at present [14]. In our studies, we found that both $CD44^+CD166^+$ lung cancer stem cells and the mice xenografts derived from $CD44^+CD166^+$ lung cancer stem cells showed abundant expression of SP-C. It showed that $CD44^+CD166^+$ lung cancer stem cells (LCSCs) were with AT-II cells' characteristic, and indicated $CD44^+CD166^+$ lung cancer stem cells might be originated from malignantly transformed AT-II cells. Therefore, the identification of $CD44^+CD166^+$ lung cancer stem cells provided an experimental basis for studying how SLC34A2 played in the malignant transformation of AT-II cells into NSCLCs.

Then, we tested the role of SLC34A2 in $CD44^+CD166^+$ lung cancer stem cells (LCSCs). Firstly, the higher expression of SLC34A2 was found in LCSCs than in their corresponding NSCLC cell lines and differentiated cells, which preliminarily indicated SLC34A2 might be related to maintaining stemness of LCSCs. Furtherly, reduced abilities of sphere formation and colony formation *in vitro*, as well as tumorigenesis *in vivo* were found in LCSCs after *SLC34A2* was interfered. Moreover, the mRNA expression levels of genes related to stemness, such as *Bmi-1, CXCR4, Nanog, β-catenin*, and *Sox2* were significantly decreased in A549-LCSCs after *SLC34A2* was interfered, and the mRNA expression level of *Sox2* was reduced in H460-LCSCs after *SLC34A2* was interfered. The evidence above demonstrated that *SLC34A2* could maintain stemness of LCSCs.

Next, we studied how *SLC34A2* maintained stemness of LCSCs. What caught our attention was that the reduced mRNA expression level of *Sox2* was found both in A549-LCSCs and H460-LCSCs after *SLC34A2* was interfered. Several reports identified that *Sox2* played a critical role in maintaining the self-renewal of embryonic stem cells and the development of lung, as well as side population (SP) cells in NSCLCs [33–35]. To our knowledge, *SLC34A2* played an important role in embryonic stem cells and the development of the lung [22, 36]. And in this study, we found *SLC34A2* had an important action in maintain stemness of LCSCs. Based on both *Sox2* and *SLC34A2* had alike functions in the development of embryonic stem cells, the lung and lung cancer stem cells, it suggested the key role of Sox2 in maintaining stemness of LCSCs by SLC34A2. Subsequently, the PI3K/AKT/STAT3 axis was detected to be downstream of SLC34A2 in LCSCs. Then we uncovered SLC34A2 could induce the transcription level of Sox2 via the PI3K/AKT/STAT3 axis. Then the enhancement effect of SLC34A2 on the binding of STAT3 to the promoter region of Sox2 was confirmed. Finally, SLC34A2 was discovered to maintain stemness of LCSCs through PI3K/AKT/STAT3/Sox2 axis *in vitro*. Besides, the connection between SLC34A2 maintaining stemness of lung cancer stem cells and the activation of PI3K/AKT/STAT3/Sox2 axis was verified *in vivo*. In the clinical specimens, we only found the $SLC34A2^{High}$ group had a higher proportion of lymphatic metastasis and higher expression of p-AKT, compared with the $SLC34A2^{Low}$ group. However,

the proportion of NSCLCs samples with high expression of PI3K, p-STAT3, and Sox2 had no difference between the two groups. It was regrettable that we could only verify the connection between SLC34A2 maintaining stemness of lung cancer stem cells and activation of AKT. The imperfect results might be due to their small sample size. In the future, we would expand the sample size for identifying it. Above all, we illustrated that *SLC34A2* could maintain stemness of LCSCs, which were with AT-II cell's characteristic. And the relative mechanism was that *SLC34A2* promoted STAT3 to bind to the promoter region of Sox2 via PI3K/AKT pathway. Furthermore, the cancer stem cell hypothesis stated that the origin of cancer was fueled by cancer stem cells (CSCs), and deregulated self-renewal signals in CSCs contributed to the initiation of cancers [3, 4]. Therefore, our results furtherly showed that the critical role and mechanism of SLC34A2 in the malignant transformation of AT-II cells into NSCLCs.

It was worth noting that in our *in vivo* studies, we injected $CD44^+CD166^+$ lung cancer stem cells (LCSCs) into the back of nude mice. Then the xenografts derived from LCSCs were found to show positive coexpression of CD44 and CD166. It suggested that the NSCLCs cells, which were derived from $CD44^+CD166^+$ lung cancer stem cells, had positive coexpression of CD44 and CD166. As we mentioned previously, 18 NSCLCs tissues with downregulated expression of SLC34A2 showed positive expression of SP-C. It indicated that the 18 NSCLCs tissues might be originated from AT-II cells and downregulated expression of SLC34A2 might lead to malignant transformation of AT-II cells. Moreover, the coexpression of CD44 and CD166 was also detected in these 18 NSCLCs tissues. Like our *in vivo* studies, the results in clinic samples suggested these 18 NSCLCs tissues might be originated from $CD44^+CD166^+$ lung cancer stem cells. The results above furtherly prompted that downregulated expression of SLC34A2 might lead AT-II cells to malignantly transform into $CD44^+CD166^+$ lung cancer stem cells (LCSCs).

Additionally, our *in vitro* findings showed the connection between the downregulated expression of SLC34A2 and the differentiation of LCSCs. The genetic alterations might make stem cells to malignantly transform into cancer stem cells, then cancer stem cells were differentiated into cancer cells [5]. We found that the expression of SLC34A2 was decreased after LCSCs were differentiated. And the morphological type of the differentiated cells derived from LCSCs was similar to their origin NSCLC cells. Meanwhile, lentiviral-si-SLC34A2 reduced the protein expression of differentiation markers CK8 and CK18 in LCSCs. These *in vitro* results preliminarily indicated that downregulated expression of SLC34A2 might induce LCSCs to differentiate into NSCLCs. However, until now we did not know why the expression of SLC34A2 was decreased during the differentiation of LCSCs and how SLC34A2 played in the differentiation of LCSCs. In future research, we will focus on this part.

However, what made us confused was that, there were significant differences in the expression and function of SLC34A2 between $CD44^+CD166^+$ lung cancer stem cells and NSCLCs cells. In our previous study, we found that SLC34A2 was significantly down-regulated both in NSCLC tissues and cell lines compared to normal cells or adjacent tissues. And overexpressed SLC34A2 could diminish cell growth, motility and invasiveness of some NSCLC cell lines in vitro and significantly attenuate tumorigenicity in vivo [25, 26]. And SLC34A2 has been indicated as a tumor-repressive role. However, in this work, we found

the significant differential expression and function of SLC34A2 between $CD44^+CD166^+$ lung cancer stem cells and NSCLCs cells. Paradoxically, we found both the mRNA and protein expression levels of SLC34A2 in LCSCs were relatively higher than that in NSCLCs cell lines. Furthermore, SLC34A2 could maintain stemness of $CD44^+CD166^+$ lung cancer stem cells, such as maintain the ability of sphere formation, colony formation in vitro and tumorigenesis potential in vivo and so on. While knockdown *SLC34A2* in LCSCs with si-RNA made the weaker abilities of sphere formation, colony formation *in vitro* and tumorigenesis potential *in vivo*. The contradictory expression and function of SLC34A2 between $CD44^+CD166^+$ lung cancer stem cells and NSCLCs cells was similar to the pattern of Bone Morphogenetic Proteins (BMPs) in glioma stem cells and gliomas. BMPs had been demonstrated to drive astrocytic differentiation in CSCs, which indicated a tumor-repressive role and were proposed as anti-CSC therapies to induce differentiation [37]. But several groups have demonstrated elevated BMP expression, particularly BMP2, was detected in human gliomas. Later the researchers demonstrated the reason that CSCs could express Gremlin1, a BMP antagonist, to specifically blocks differentiation effects of BMPs [37], which gave a reasonable explanation that contradictory expression of BMPs in glioma stem cells and gliomas. Although we first found the significant differential expression and function of SLC34A2 between $CD44^+CD166^+$ lung cancer stem cells and NSCLCs cells, however, we were not clear about the reason for the contradictory result now. We are interested in the seemingly contradictory result and would carry on exploration to this question in the future.

Above all, our studies demonstrated SLC34A2 played an important role in the malignant transformation of AT-II cells into NSCLCs. The downregulated expression of SLC34A2 was found in NSCLCs cells than the AT-II cells in clinic samples. Then SLC34A2 was detected to maintain the stemness of $CD44^+CD166^+$ lung cancer stem cells (LCSCs), which were with AT-II cells' characteristic. Finally, the relative mechanism was uncovered that SLC34A2 induced transcription of Sox2 by promoting the binding of STAT3 to the promoter region of Sox2 via PI3K/AKT pathway. Additionally, we suggested the downregulated expression of SLC34A2 performed a vital role in the process that AT-II cells malignantly transformed into LCSCs, and then differentiated into NSCLCs (Additional File 11). It showed the critical role and mechanism of SLC34A2 in the malignant transformation of AT-II cells and initiation of NSCLCs. Also it provided a new insight for exploring how AT-II cells malignantly transform into NSCLCs.

Conclusions

Above all, we illustrated the expression of SLC34A2 was downregulated in NSCLCs cells compared with AT-II cells. We discovered the downregulated expression of SLC34A2 performed a vital role in the malignant transformation of AT-II cells into NSCLCs for the first time. And SLC34A2 could maintain stemness of $CD44^+CD166^+$ lung cancer stem cells, which were with AT-II cell's characteristic, via PI3K/AKT/STAT3/Sox2 axis. It had important significance in the revelation of a new mechanism for the initiation of NSCLCs.

Abbreviations

AT II：Alveolar Type II epithelium cells; Akt：Protein Kinase B; AP：Ammonium Persulfate; ATCC：American Type Culture Collection; BSA：Bovine Serum Albumin; CHIP：Chromatin Immunoprecipitation; CK8：Cytokeratin 8; CK18：Cytokeratin 18; DEPC: Diethylpyrocarbonate; DMSO: Dimethylsulphoxide; EDTA：Ethylenediaminetetraacetic acid; EGF：Epidermal Growth Factor; EMT：Epithelial Mesenchymal Transitions; FBS：Fetal Bovine Serum; FGF：Fibroblast Growth Factor; IGF：Insulin Growth Factor; IHC：Immunohistochemical; IL-6：Interleukin 6; LCSCs：$CD44^{+}CD166^{+}$ Lung Cancer Stem Cells; MTS：3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide; NF-KB：Nuclear Factor kappa-light-chain-enhancer of activated B cells; NSCLC ：Non-Small Cell Lung Cancer; OD：Optical Density; p-Akt：Phospho-Protein Kinase B; PBS：Phosphate Buffered Saline; PCR：Polymerase Chain Reaction; PI3K：Phosphatidylinositol-4,5-bisphosphate 3-kinase; p- NF-KB：Phospho-Nuclear Factor kappa-light-chain-enhancer of activated B cells; p-STAT3：Phospho-Signal transducer and activator of transcription 3; PVDF: Polyvinylidene Fluoride; QPCR：Quantitative real-time PCR; RIPA：Radioimmunoprecipitation; SDS：Sodium Dodecyl Sulfate; STAT3：Signal transducer and activator of transcription 3; Tris：2-Amino-2-(hydroxymethyl)-1,3-propanediol

Declarations

- **Ethics approval and consent to participate**

Usage of NSCLC tissues and their corresponding adjacent tissues in our study was performed with the approval of the Medical Ethical Committee of West China Hospital, Sichuan University. Animal assays were carried out according to the National Institutes of Health ‘Guide for the Care and Use of Laboratory Animals’ and the experimental protocols in the study were approved by the animal ethical committee of West China Hospital, Sichuan University.

- **Consent for publication**

Written informed consent for publication was obtained from all participants.

- **Availability of data and material**

All data generated or analyzed during this study were included in this published article and its supplementary information files. And all data supporting the findings of this study are available from the corresponding author on reasonable request. TCGA database analysis was performed by querying the website: cBioPortal for Cancer Genomics (http://www.cbioportal.org/)

- **Competing interests**

The authors declare that they have no competing interests.

- **Funding**

This study was supported by Sichuan Provincial Department of Science and Technology | Applied Basic Research Program of Sichuan Province No. 2018JY0134).

- **Authors' contributions**

Yan Yang: collection and/or assembly of data, data analysis and interpretation, manuscript writing, final approval of the manuscript; Wen Zhu: conception and design, financial support, administrative support, provision of study material, data analysis and interpretation, manuscript writing, final approval of the manuscript; Qiang Pu, Hu Liao, Yue Yuan, Xueting Hu, Xingmei Luo, Qianqian Jiang, Xiaolan Su: collection and/or assembly of data, data analysis and interpretation, technical support; Qiang Pu and Hu Liao : provision of study material or patients.

- **Acknowledgments**

The authors thank Qiaorong Huang, Tingting Guo, Jie Su and Yi Zhang for sorting and identification of LCSCs, immunohistochemistry and animal experimental assistance.

References

1. Gridelli C, Rossi A, Carbone DP, Guarize J, Karachaliou N, Mok T, Petrella F, Spaggiari L, Rosell R: **Non-small-cell lung cancer.***Nat Rev Dis Primers* 2015, **1:**1-16.
2. Siegel RL, Miller KD, Jemal A: **Cancer statistics, 2018.***CA Cancer J Clin* 2018, **68:**7-30.
3. Batlle E, Clevers H: **Cancer stem cells revisited.***Nat Med* 2017, **23:**1124-1134.
4. Borah A, Raveendran S, Rochani A, Maekawa T, Kumar DS: **Targeting self-renewal pathways in cancer stem cells: clinical implications for cancer therapy.***Oncogenesis* 2015, **4:**e177.
5. Heng WS, Gosens R, Kruyt FAE: **Lung cancer stem cells: origin, features, maintenance mechanisms and therapeutic targeting.***Biochemical pharmacology* 2019, **160:**121-133.
6. Ain Zubaidah Ayob, Thamil Selvee Ramasamy: **Cancer stem cells as key drivers of tumour progression.***Journal of Biomedical Science,* 2018, 25: 20.
7. Lundin A, Driscoll B: **Lung cancer stem cells: progress and prospects.***Cancer Lett* 2013, **338:**89-93.
8. Lau AN, Curtis SJ, Fillmore CM, Rowbotham SP, Mohseni M, Wagner DE, Beede AM, Montoro DT, Sinkevicius KW, Walton ZE, et al: **Tumor-propagating cells and Yap/Taz activity contribute to lung tumor progression and metastasis.***EMBO J* 2014, **33:**468-481.
9. Kim CF: **Intersections of lung progenitor cells, lung disease and lung cancer.***Eur Respir Rev* 2017, **26:**1-6.
10. Leon G, MacDonagh L, Finn SP, Cuffe S, Barr MP: **Cancer stem cells in drug resistant lung cancer: Targeting cell surface markers and signaling pathways.***Pharmacol Ther* 2016, **158:**71-90.
11. Desai TJ, Brownfield DG, Krasnow MA: **Alveolar progenitor and stem cells in lung development, renewal and cancer.***Nature* 2014, **507:**190-194

12. Beers MF, Moodley Y: **When Is an Alveolar Type 2 Cell an Alveolar Type 2 Cell? A Conundrum for Lung Stem Cell Biology and Regenerative Medicine.***Am J Respir Cell Mol Biol* 2017, **57:**18-27.
13. Guillot L, Nathan N, Tabary O, Thouvenin G, Le Rouzic P, Corvol H, Amselem S, Clement A: **Alveolar epithelial cells: master regulators of lung homeostasis.***Int J Biochem Cell Biol* 2013, **45:**2568-2573.
14. Treutlein B, Brownfield DG, Wu AR, Neff NF, Mantalas GL, Espinoza FH, Desai TJ, Krasnow MA, Quake SR: **Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq.***Nature* 2014, **509:**371-375..
15. Ten Have-Opbroek AA, Benfield JR, van Krieken JH, Dijkman JH: **The alveolar type II cell is a pluripotential stem cell in the genesis of human adenocarcinomas and squamous cell carcinomas.***Histol Histopathol* 1997, **12:**319-336.
16. Kitinya JN, Sueishi K, Tanaka K, Katsuda Y: **Immunoreactivity of surfactant-apoprotein in adenocarcinomas, large cell and small cell carcinomas of the lung.***Acta Pathol Jpn* 1986, **36:**1271-1278.
17. Xu X, Rock JR, Lu Y, Futtner C, Schwab B, Guinney J, Hogan BL, Onaitis MW: **Evidence for type II cells as cells of origin of K-Ras-induced distal lung adenocarcinoma.***Proc Natl Acad Sci U S A* 2012, **109:**4910-4915.
18. Mainardi S, Mijimolle N, Francoz S, Vicente-Duenas C, Sanchez-Garcia I, Barbacid M: **Identification of cancer initiating cells in K-Ras driven lung adenocarcinoma.***Proc Natl Acad Sci U S A* 2014, **111:**255-260.
19. Xu X, Huang L, Futtner C, Schwab B, Rampersad RR, Lu Y, Sporn TA, Hogan BL, Onaitis MW: **The cell of origin and subtype of K-Ras-induced lung tumors are modified by Notch and Sox2.***Genes Dev* 2014, **28:**1929-1939.
20. Murer H, Forster I, Biber J: **The sodium phosphate cotransporter family SLC34.***Pflugers Arch* 2004, **447:**763-767.
21. Zheng X, Kammerer CM, Cox LA, Morrison A, Turner ST, Ferrell RE: **Association of SLC34A2 variation and sodium-lithium countertransport activity in humans and baboons.***Am J Hypertens* 2009, **22:**288-293.
22. Xu H, Bai L, Collins JF, Ghishan FK: **Molecular cloning, functional characterization, tissue distribution, and chromosomal localization of a human, small intestinal sodium-phosphate (Na+-Pi) transporter (SLC34A2).***Genomics* 1999, **62:**281-284.
23. Jacob A, Morley M, Hawkins F, McCauley KB, Jean JC, Heins H, Na CL, Weaver TE, Vedaie M, Hurley K, et al: **Differentiation of Human Pluripotent Stem Cells into Functional Lung Alveolar Epithelial Cells.***Cell Stem Cell* 2017, **21:**472-488.
24. Kopantzev EP, Monastyrskaya GS, Vinogradova TV, Zinovyeva MV, Kostina MB, Filyukova OB, Tonevitsky AG, Sukhikh GT, Sverdlov ED: **Differences in gene expression levels between early and later stages of human lung development are opposite to those between normal lung tissue and non-small lung cell carcinoma.***Lung Cancer* 2008, **62:**23-34.

25. Zhang X, Ke X, Pu Q, Yuan Y, Yang W, Luo X, Jiang Q, Hu X, Gong Y, Tang K, et al: **MicroRNA-410 acts as oncogene in NSCLC through downregulating SLC34A2 via activating Wnt/beta-catenin pathway.***Oncotarget* 2016, **7:**14569-14585.
26. Wang Y, Yang W, Pu Q, Yang Y, Ye S, Ma Q, Ren J, Cao Z, Zhong G, Zhang X, et al: **The effects and mechanisms of SLC34A2 in tumorigenesis and progression of human non-small cell lung cancer.***J Biomed Sci* 2015, **22:**52.
27. Mitsuyoshi Hashimoto, Dong-Yu Wang, Takaharu Kamo, Yue Zhu, Toshifumi Tsujiuchi, Yoichi Konishi, Masamitsu Tanaka, Haruhiko Sugimura: **Isolation and Localization of Type IIb Na/Pi Cotransporter in the Developing Rat Lung.** *American Journal of Pathology,* 2000, **157:**21-27.
28. Shukla S, Khan S, Sinha S, Meeran SM: **Lung Cancer Stem Cells: An Epigenetic Perspective.***Curr Cancer Drug Targets* 2018, **18:**16-31.
29. Zakaria N, Satar NA, Abu Halim NH, Ngalim SH, Yusoff NM, Lin J, Yahaya BH: **Targeting Lung Cancer Stem Cells: Research and Clinical Impacts.***Front Oncol* 2017, **7:**80.
30. Chen Q, Suresh Kumar V, Finn J, Jiang D, Liang J, Zhao YY, Liu Y: **CD44(high) alveolar type II cells show stem cell properties during steady-state alveolar homeostasis.***Am J Physiol Lung Cell Mol Physiol* 2017, **313:**L41-L51.
31. Soh BS, Zheng D, Li Yeo JS, Yang HH, Ng SY, Wong LH, Zhang W, Li P, Nichane M, Asmat A, et al: **CD166(pos) subpopulation from differentiated human ES and iPS cells support repair of acute lung injury.***Mol Ther* 2012, **20:**2335-2346.
32. Zhao D, Pan C, Sun J, Gilbert C, Drews-Elger K, Azzam DJ, Picon-Ruiz M, Kim M, Ullmer W, El-Ashry D, et al: **VEGF drives cancer-initiating stem cells through VEGFR-2/Stat3 signaling to upregulate Myc and Sox2.***Oncogene* 2015, **34:**3107-3119.
33. Ben-Porath I, Thomson MW, Carey VJ, Ge R, Bell GW, Regev A, Weinberg RA: **An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors.***Nat Genet* 2008, **40:**499-507.
34. Tompakins DH, Besnard V, Lange AW, Wert SE, Keiser AR, Smith AN, Lang R, Whitsett JA: **Sox2 is required for maintenance and differentiation of bronchiolar Clara, ciliated, and goblet cells.***PLoS One* 2009, **4:**e8248.
35. Danopoulos S, Alonso I, Thornton ME, Grubbs BH, Bellusci S, Warburton D, Al Alam D: **Human lung branching morphogenesis is orchestrated by the spatiotemporal distribution of ACTA2, SOX2, and SOX9.***Am J Physiol Lung Cell Mol Physiol* 2018, **314:**L144-L149.
36. Singh S, Trevino J, Bora-Singhal N, Coppola D, Haura E, Altiok S, Chellappan SP: **EGFR/Src/Akt signaling modulates Sox2 expression and self-renewal of stem-like side-population cells in non-small cell lung cancer.***Mol Cancer* 2012, **11:**73.
37. Shibasaki Y, Etoh N, Hayasaka M, Takahashi M, Kakitani M, Yamashita T, Tomizuka K, Hanaoka K: **Targeted deletion of the tybe IIb Na+-dependent Pi-co-transporter, NaPi-IIb, results in early embryonic lethality.***Biochemical and biophysical research communications* 2009, **381:**482-486.

38. Kenneth Yan, Qiulian Wu, Diana H. Yan, Christine H. Lee, Nasiha Rahim, Isabel Tritschler,Jennifer DeVecchio, Matthew F. Kalady, Anita B. Hjelmeland, and Jeremy N. Rich: **Glioma cancer stem cells secrete Gremlin1 to promote their maintenance within the tumor hierarchy**. *Genes & Development* 2014, **28**:1085–1100.

Figures

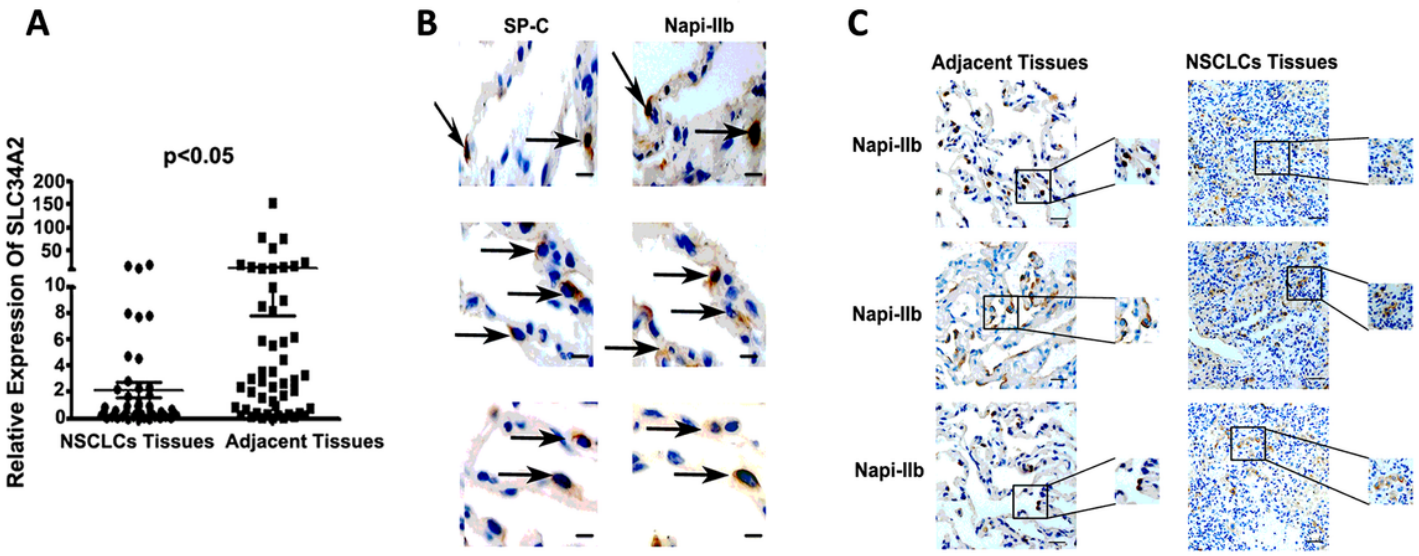

Figure 1

Downregulated expression of SLC34A2 was involved in malignant transformation of AT-II cells in clinic samples (A) The relative mRNA expression level of SLC34A2 in 57 collected NSCLC tissues and their adjacent tissues. (B)Immunohistochemical staining of SP-C and Napi-IIb in collected NSCLC adjacent tissues (400x, scale bar 50 μm). (C) Representative IHC images of the cell with positive expression of Napi-IIb in NSCLCs tissue section and adjacent tissue section (200x, scale bar 100μm).

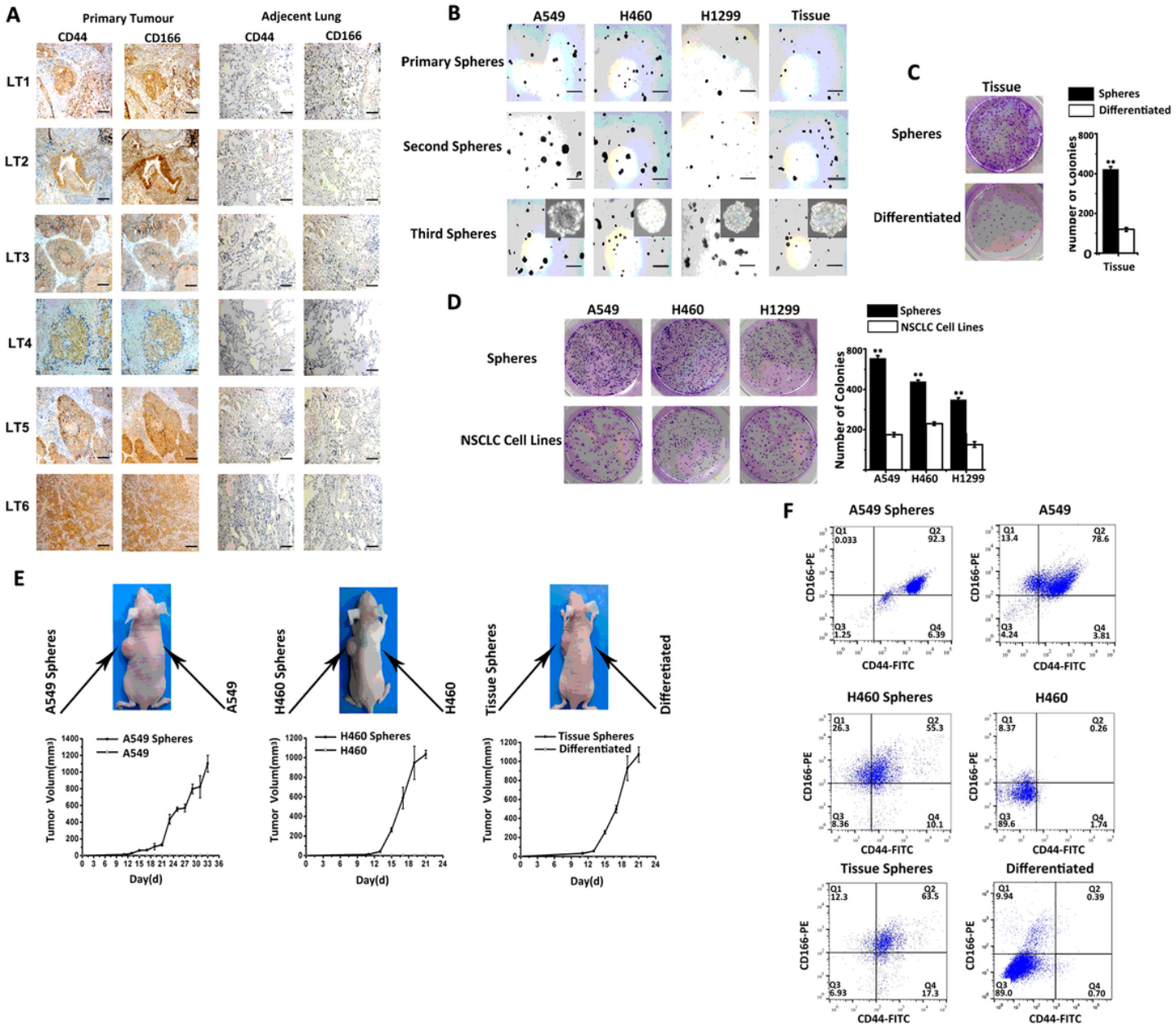

Figure 2

The Expressions of CD44 and CD166 Were Evaluated in NSCLCs Samples and Spheres from NSCLCs Cell Lines and NSCLCs Tissue. (A) Representative IHC images of CD44 and CD166 in NSCLCs tissue sections and their corresponding adjacent tissue sections (200x, scale bar 100 μm). (B)Examples of spheres generated from NSCLC cell lines (A549, H460, and H1299) and tissue with 100× original magnification. Primary spheres: spheres directly generated from NSCLC tissue and NSCLC cell lines. Second and third spheres: the second and third passage spheres from primary spheres (Scale bar, 200 μm). (C) The ability of colony formation of spheres generated from NSCLC tissue cultured in serum-free medium and their differentiated cells cultured in serum medium. (D) The ability of colony formation of spheres generated from NSCLC cell lines cultured in serum-free medium and their corresponding NSCLC cell lines cultured in serum medium. (E) The ability of tumor initiation of spheres, NSCLC cell lines and differentiated cells

from NSCLC tissue in nude mice (n=4). (F) FACS analysis of CD44+CD166+ expression in spheres and their corresponding NSCLC cell lines or differentiated cells. Note: Differentiated: The differentiated cells which were obtained from the spheres generated from NSCLC tissue by culturing in DMEM/F12/10% FBS medium for 1 week.

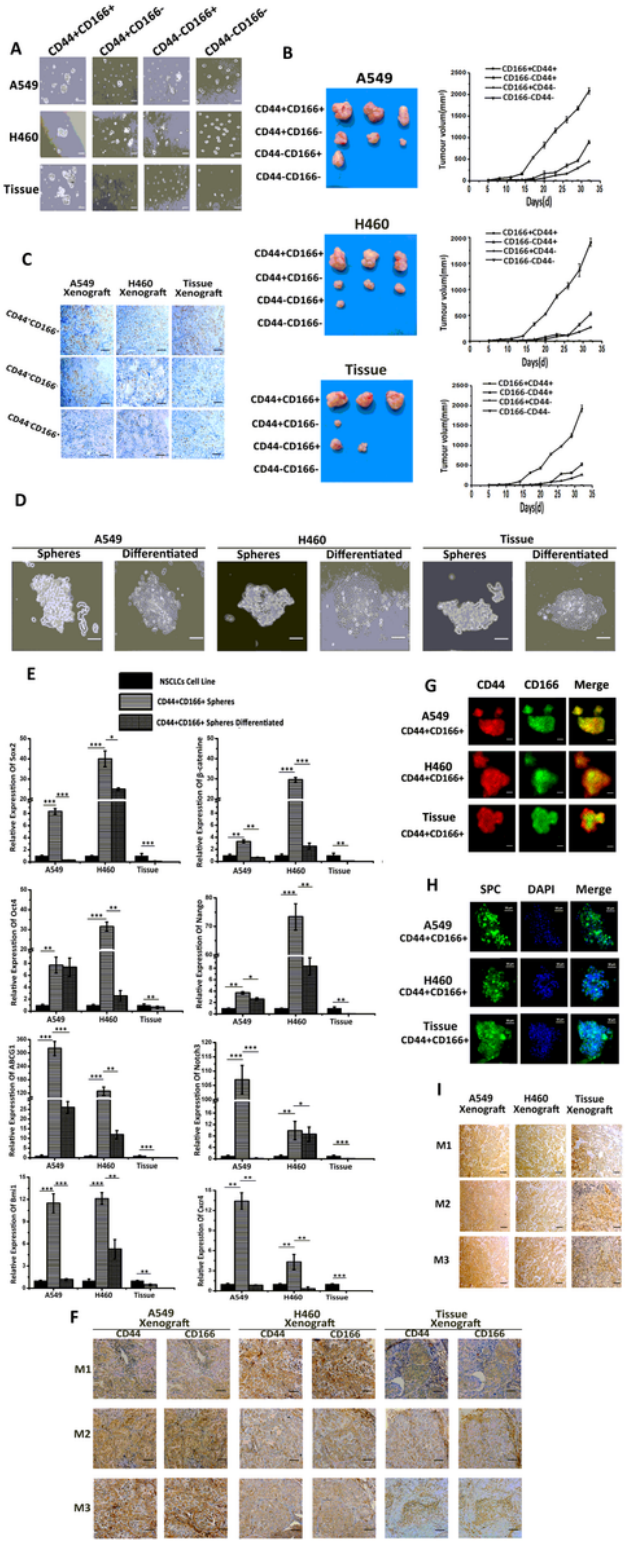

Figure 3

The CD44+CD166+ fraction were identified to have stemness and AT-II cells' characteristic. (A) The ability of sphere formation of CD44+CD166+, CD44-CD166+ CD44+CD166- and CD44-CD166- fractions in serum-free medium with 100× original magnification. Scale bar, 200 μm. (B) The potential of tumor initiation of CD44+CD166+, CD44-CD166+, CD44+CD166- and CD44-CD166- fractions by subcutaneous injection and representative tumor-growth curves of xenografts derived from different cell fractions. (C) Ki-67 analysis of corresponding xenografts obtained after the injection of spheres with 200× original magnification. Scale bar, 100 μm.(D) Images of CD44+CD166+ spheres grown as undifferentiated cells (spheres) or under differentiative conditions (differentiated) with 200× original magnification. Scale bar, 100 μm. (E) Expression of genes involved in stemness and drug resistance in CD44+CD166+ sphere cells, differentiated cells derived from CD44+CD166+ spheres and their corresponding NSCLC cell lines. Note: Data represent the mean±s.d. *P<0.05,**P<0.01. (F) Expression of CD44 and CD166 in corresponding xenografts obtained from CD44+CD166+ spheres by IHC with 200× original magnification. Scale bar, 100 μm. M1: mouse 1. M2: mouse 2. M3: mouse 3. (G) Expression of CD44 and CD166 in FACS-purified fractions of CD44+CD166+ sphere progenies by immunofluorescence staining with 400× original magnification. Scale bar, 50 μm. (H)Expression of SP-C in FACS-purified fractions of CD44+CD166+ sphere progenies by immunofluorescence staining with 400× original magnification. Scale bar, 50 μm. (I) Expression of SP-C in corresponding xenografts obtained after the injection of sphere cells with 200× original magnification. M1: mouse 1. M2: mouse 2. M3: mouse 3. Scale bar, 100 μm.

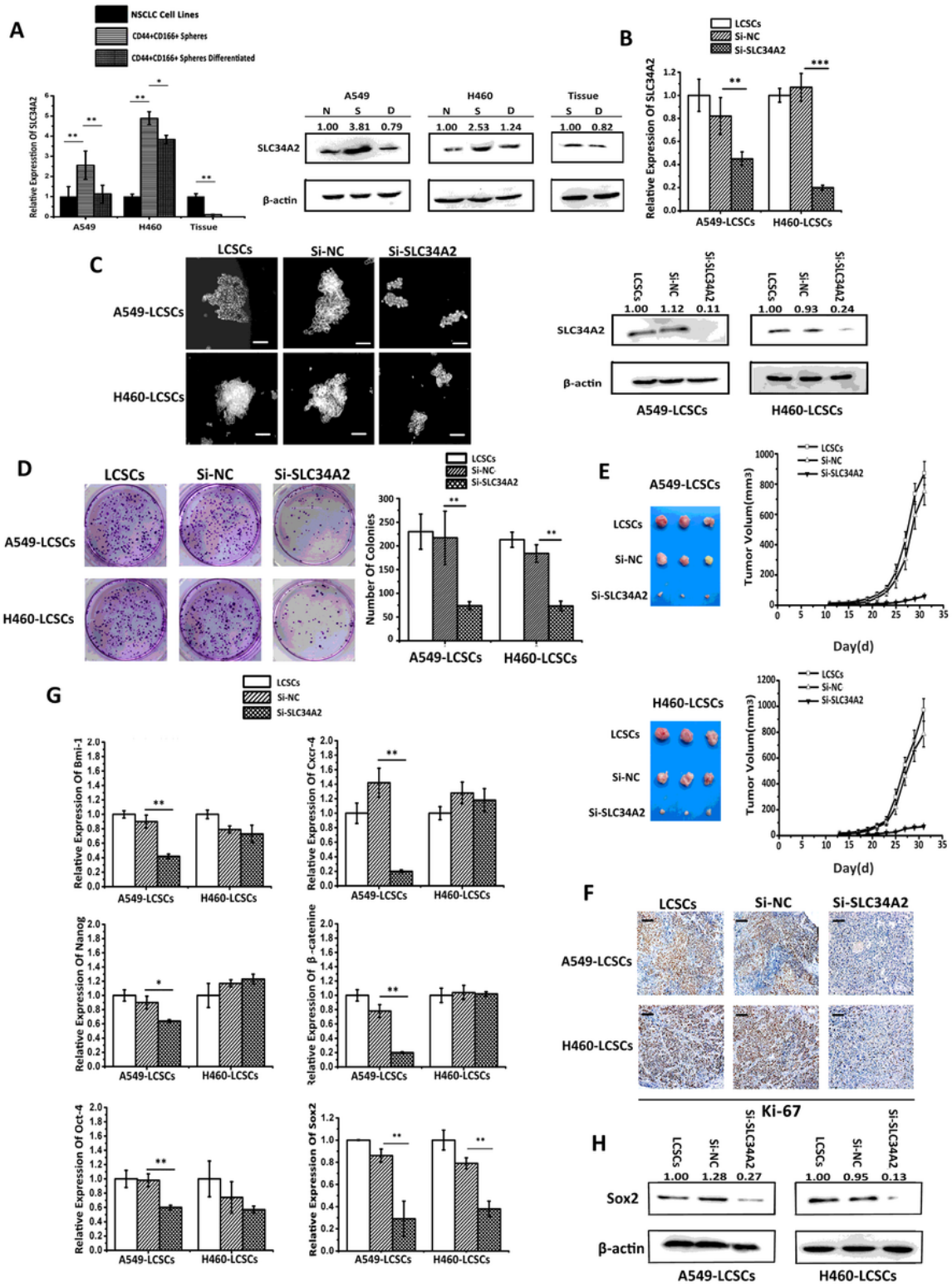

Figure 4

SLC34A2 was found to maintain stem cell-like characteristics of LCSCs in vitro and in vivo. (A) Relative mRNA and protein expression level of SLC34A2 in LCSCs, differentiated cells derived from LCSCs and their corresponding NSCLC cell lines. (B) Relative mRNA and protein expression level of SLC34A2 in LCSCs, LCSCs transfected with si-SLC34A2 (si-SLC34A2) and transfected with si-NC (si-NC). (C) The ability of sphere formation of LCSCs, LCSCs transfected with si-SLC34A2 (si-SLC34A2) and transfected

with si-NC (si-NC) with 400× original magnification. Scale bar, 50 μm. (D) The ability of colony formation of LCSCs, LCSCs transfected with si-SLC34A2 (si-SLC34A2) and transfected with si-NC (si-NC). Note: Data represent the mean±s.d. *P<0.05,**P<0.01. (E) The potential of tumor initiation of LCSCs, LCSCs transfected with si-SLC34A2 (si-SLC34A2) and transfected with si-NC (si-NC), and tumor-growth curves of xenografts derived from different cell fractions. (F) Ki-67 analysis of corresponding xenografts derived from LCSCs, LCSCs transfected with si-SLC34A2 (si-SLC34A2) and transfected with si-NC (si-NC) with 200 × original magnification. Scale bar, 100 μm. (G) The mRNA expression of genes related to stemness in LCSCs, LCSCs transfected with si-SLC34A2 (si-SLC34A2) and transfected with si-NC (si-NC). (H) The protein expression of Sox2 in LCSCs, LCSCs transfected with si-SLC34A2 (si-SLC34A2) and transfected with si-NC (si-NC). Note: LCSCs was short for CD44+CD166+ lung cancer stem cells. si-NC was short for LCSCs transfected with si-NC and si-SLC34A2 was short for LCSCs transfected with si-SLC34A2.

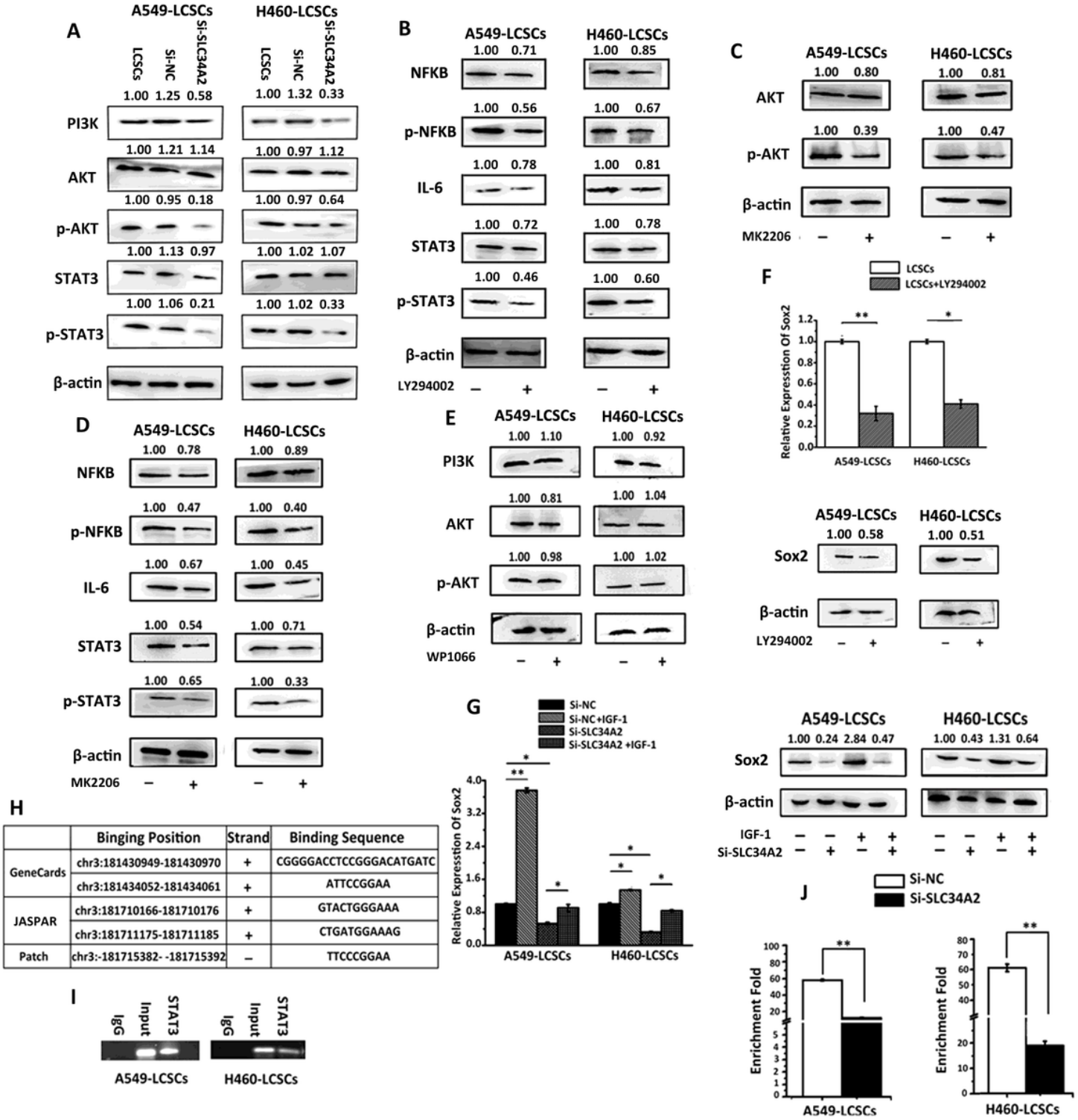

Figure 5

SLC34A2 was discovered to promote STAT3 to bind to the promoter region of Sox2 via PI3K/AKT pathway. (A) The protein expression level of key factors in PI3K/AKT and STAT3 pathways in LCSCs, LCSCs transfected with si-SLC34A2 (si-SLC34A2) and transfected with si-NC (si-NC). (B) The expression of key proteins in the STAT3 pathway in LCSCs after LY294002 treatment. (C) The expression of phosphorylation AKT in LCSCs after MK2206 treatment. (D) The expression of key proteins in the STAT3 pathway in LCSCs after MK2206 treatment. (E) The expression of key proteins in the PI3K/AKT pathway

in LCSCs after WP1066 treatment. (F) The mRNA and protein expression level of Sox2 in LCSCs after LY294002 treatment. (G) The mRNA and protein expression level of Sox2 in si–SLC34A2-LCSCs after IGF-1 treatment. (H) Binding sites of STAT3 in the promoter region of Sox2 were predicted by bioinformatics software Genecards, JASPAR and Patch. (I) The amplified band of the binding sequence of STAT3 to the promoter region of Sox2 in LCSCs. (J) The binding probability of STAT3 to the promoter region of Sox2 in si–SLC34A2-LCSCs and si–NC-LCSCs. Note: LCSCs was short for CD44+CD166+ lung cancer stem cells. si-NC was short for LCSCs transfected with si-NC and si-SLC34A2 was short for LCSCs transfected with si-SLC34A2.

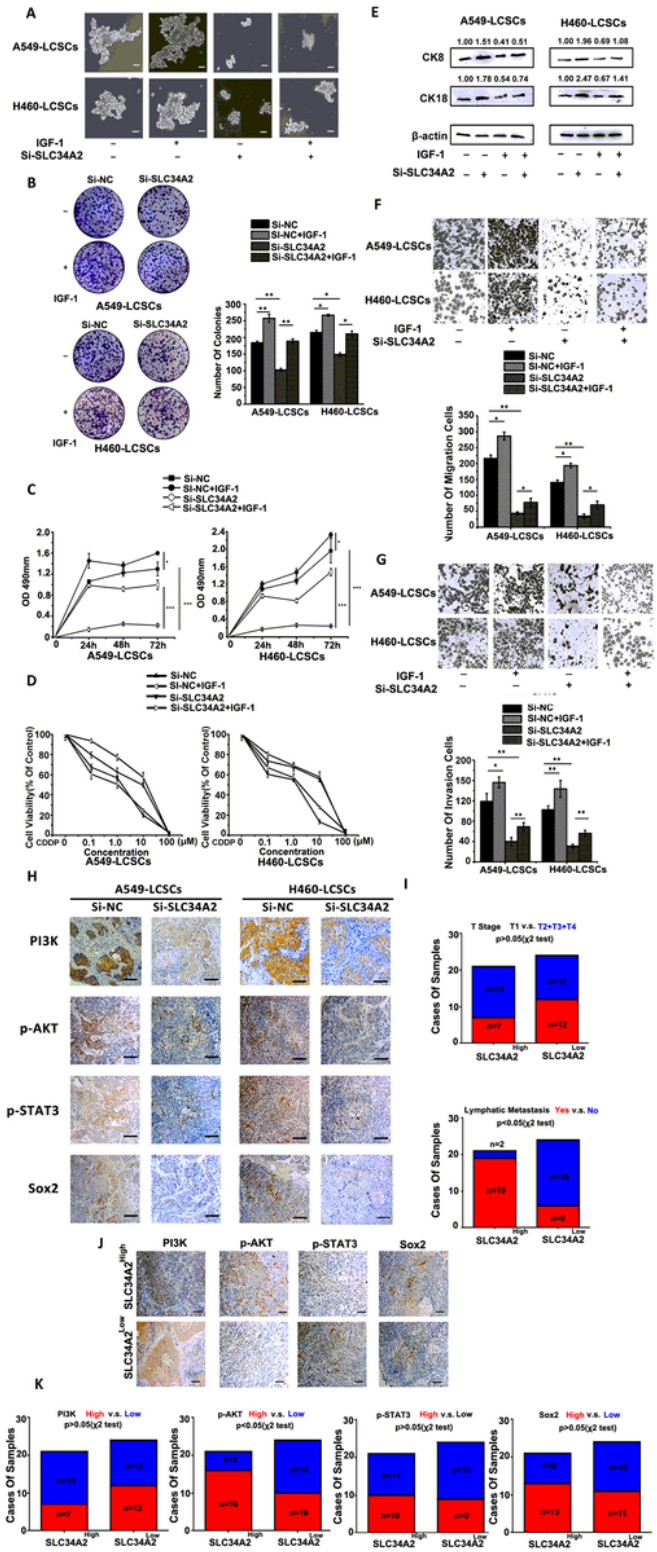

Figure 6

The mechanism of SLC34A2 maintaining stemness of LCSCs was uncovered to be through PI3K/AKT/STAT3/Sox2 axis. (A) The ability of sphere formation of si-SLC34A2-LCSCs after IGF-1 treatment with 200 × magnifications. Scale bar, 100 μm. (B) The ability of colony formation of si-SLC34A2-LCSCs after IGF-1 treatment. (C) The growth rate of si-SLC34A2-LCSCs after IGF-1 treatment by MTS assay. (D) The viability against cisplatin of si-SLC34A2-LCSCs after IGF-1 treatment. (E) The

expression of CK8 and CK18 in si-SLC34A2-LCSCs after IGF-1 treatment. (F-G) The ability of migration and invasion of si-SLC34A2-LCSCs after IGF-1 treatment with 200 × magnifications. Scale bar, 100 μm. (H) Expression of PI3K, p-AKT, p-STAT3, and Sox2 in xenografts derived from si-NC and si-SLC34A2. (200 x, Scale bar, 100 μm). (I) The difference in the proportion of NSCLCs samples with T stage or lymphatic metastasis between the group with the relatively higher expression of SLC34A2 (SLC34A2High) and relatively lower expression of SLC34A2 (SLC34A2Low). (J)Representative IHC images of PI3K, p-AKT, p-STAT3, and Sox2 in the group with relatively higher expression of SLC34A2 and relatively lower expression of SLC34A2 (200x, scale bar 100 μm). (K) The difference in the proportion of NSCLCs samples with high expression of PI3K, p-AKT, p-STAT3 and Sox2 between the group with relative higher expression of SLC34A2 and relative lower expression of SLC34A2. Note: LCSCs was short for CD44+CD166+ lung cancer stem cells. si-NC was short for LCSCs transfected with si-NC and si-SLC34A2 was short for LCSCs transfected with si-SLC34A2.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AdditionalFile1MaterialsandMethods.docx
- AdditionalFile2TableS1.docx
- AdditionalFile3TableS2.docx
- AdditionalFile4TableS3.docx
- AdditionalFile5TableS4.docx
- AdditionalFile6FigureS1.tif
- AdditionalFile7TableS5.docx
- AdditionalFile8FigureS2.tif
- AdditionalFile9FigureS3.tif
- AdditionalFile10TheCancerGenomeAtlasTCGAanalysis.docx
- AdditionalFile11FigureS5.tif