Article:
Town, J, Pais, H, Harrison, S et al. (5 more authors) (2016) Exploring the surfaceome of Ewing sarcoma identifies a new and unique therapeutic target. Proceedings of the National Academy of Sciences, 113 (13).

https://doi.org/10.1073/pnas.1521251113

Exploring the surfaceome of Ewing sarcoma identifies a novel and unique therapeutic target

Jennifer Town[1], Helio Pais[1], Sally Harrison[2, 4], Lucy F. Stead[2, 4], Carole Bataille[3], Wilawan Bunjobpol[1], Jing Zhang[1] and Terence H. Rabbitts[1]

[1] Weatherall Institute for Molecular Medicine, MRC Molecular Haematology Unit, University of Oxford, John Radcliffe Hospital, Oxford, OX3 9DS, UK

[2] Leeds Institute of Molecular Medicine, St. James's Hospital, Leeds, LS9 7TF, UK

[3] Chemistry Research Labs, Dept. of Chemistry, Mansfield Road, Oxford OX1 3AT, UK

[4] Current address: Leeds Institute of Cancer & Pathology, St. James's Hospital, Leeds, LS9 7TF, UK

Key words: Ewing sarcoma; cell surface; surfaceome; therapy; antibody; LINGO1; cancer

Running title: Ewing sarcoma surfaceome & LINGO1

Significance

By investigating cell surface proteins of Ewing sarcoma we have identified an antigen that is uniquely expressed on these tumour cells compared with mesenchymal stem cells. This protein acts as a target for antibody drug conjugates that are internalized and can kill these tumour cells, presaging translating to clinical use in treating Ewing sarcoma, especially metastatic disease.

Abstract

The cell surface proteome of tumours mediates the interface between the transformed cells and the general micro-environment including interactions with stromal cells in the tumour niche and immune cells such as T cells. In addition, the cell surface proteome of individual cancers defines biomarkers for that tumour type and potential proteins that can be the target of antibody-mediated therapy. We have used next generation deep RNA sequencing (RNA-seq) coupled to an in-house database of genes encoding cell surface proteins (herein referred to as the surfaceome) as a tool to define a cell surface proteome of Ewing sarcoma compared with progenitor mesenchymal stem cells. This subtractive RNA-seq analysis revealed a specific surfaceome of Ewing and showed unexpectedly that the leucine-rich repeat and immunoglobulin domain protein LINGO1 is expressed on over 90% of Ewing sarcoma tumours, but not expressed in any other somatic tissue apart from the brain. We found that the LINGO1 protein acts as a gateway protein internalizing into the tumour cells when engaged by antibody and can carry antibody conjugated with drugs to kill Ewing sarcoma cells. Therefore, LINGO1 is a novel, unique and specific biomarker and drug target for the treatment of Ewing sarcoma.

/body

Introduction

Targeted cancer treatment options rely on the identification of specific target proteins that allow the differentiation between normal and malignant cells. Monoclonal antibodies that selectively bind to such target proteins have been successfully utilized in the clinic. Cell surface proteins are excellent targets for antibody-based therapeutics due to their accessibility. Mechanisms by which antibodies can induce tumour cell killing include Antibody-Dependent Cell mediated Cytotoxicity (ADCC) and specific delivery of a cytotoxic payload to tumour cells using antibody-drug conjugates (ADC) (1). However, there are few cancer-specific cell surface proteins that can be invoked for antibody targeting. Of the predicted number of about 21,000 human genes, approximately 4700 are predicted to be membrane associated. Methods are required to filter this information and allow further prediction of cell surface molecules. The analysis of whole cellular transcriptomes by next generation deep RNA sequencing (RNA-seq) is a new method for target discovery, which can be utilized as a surrogate tool for the analysis of the proteome including the entirety of cell surface proteins, called the surfaceome (2, 3).

The Ewing sarcoma family of tumours (ESFT) are aggressive bone and soft tissue tumours with a high propensity to metastasise. Ewing sarcoma is the second most common bone tumour of children and adolescents with the mean age of diagnosis being 15 years of age.(4, 5). The current standard of care treatment is multimodal treatment including systemic chemotherapy with either radiation or surgery often with limb amputation in patients with local recurrence (6, 7). However, despite aggressive treatment the 5-year survival rate is 60-70% for localized disease and drops sharply to only 30% when the cancer metastasises (4, 8). There is therefore a need for novel targeted therapies for these cancers, that will overcome the limitations of the current

treatment regimens, namely the severe side effects and very limited effectiveness for metastasized disease.

Ewing sarcoma arises as a consequence of balanced chromosomal translocations leading to an in frame fusion of the EWSR1 gene with a member of the ETS family of genes, principally resulting fusion protein is EWS-FLI1, which acts as an aberrant transcription factor and induces global changes in gene expression that are essential for malignant transformation and tumour formation (4). The most likely cells of origin of Ewing sarcoma are mesenchymal stem cells (MSCs) since these cells are permissive for EWS-FLI1, which is toxic for many cell lines (9, 10). In addition, EWS-FLI1 expression in MSCs induces a gene expression profile that is highly similar to EWS (11) while EWS-FLI1 silencing in ES cell lines leads to the conversion towards a MSC expression profile (12).

To facilitate the development of a general approach to identify candidate cell surface proteins, we examined the Ewing sarcoma cell surface using a new RNA-seq surfaceome database to analyze whole transcriptomes of polyA+ RNA from three cell lines with two MSC lines (subtractive RNA-seq). Our work on the Ewing sarcoma RNA-seq is a proof-of-concept and revealed a set of candidate target proteins that are differentially expressed in the tumour cells. One of these target genes is the leucine-rich repeat and Ig domain containing protein 1 (LINGO1) first identified as a component of protein complex on brain cells (13). Our data suggest that LINGO1 is a highly specific drug target and novel biomarker of Ewing sarcoma tumours.

Results

Generation of a database of genes encoding surfaceome proteins

The use of data for genes encoding cell surface proteins (the surfaceome) (3), has increasing importance since whole transcriptome data can now be obtained by next generation deep sequencing of populations of cells and single cells (14-16). Further, complete genome sequencing have defined all the coding genes allowing classification of gene products into functional categories and into sub-cellular locations, such as nucleus, mitochondria, membrane-association, cell surface and secreted. We have developed surfaceome database based on the sources and criteria listed in SI Appendix, Table S1A. The potential surfaceome genes were classified as gold, silver or bronze (SI Appendix, Table S1B). These surfaceome database class designations were established to give degrees of confidence about the validity of each candidate surface protein encoding mRNA (for instance gold is where a protein has been shown to have surface expression) whereas the two other classes distinguish higher and lower confidence is such predictions. Accordingly, next generation deep RNA-seq data can be filtered using the surfaceome database to list genes that will encode proteins at the cell surface and, by comparing two related cell populations, candidate proteins can be identified that are differently expressed (subtractive RNA-seq).

Analysis of RNA-seq surfaceome data reveals Ewing sarcoma cell surface proteins

Possible novel surface proteins that could be therapeutic targets for Ewing sarcoma were investigated by implementing the filtering of whole transcriptome deep sequencing of RNA from three Ewing sarcoma cell lines (A673, TC-32 and TTC-466) compared with two MSC lines (5H and 4+v) (17). The complete data sets for the five RNA populations are given in Dataset S1. We compared the two whole transcriptomes with surfaceome encoding transcripts (Figure 1 and SI Appendix,

Figure S2) as well as transcription factors, cell cycle and cell signaling proteins (SI Appendix, Figures S1 and S2). The subtracted surfaceome RNA-seq data set is listed in Dataset S2. This revealed 839 mRNA transcripts that are preferentially expressed in EWS and 931 transcripts that are preferentially expressed in the MSC lines (Figure 1A). By filtering the RNA-seq data with our surfaceome database and subtracted between the two RNA-seq data sets, we found that 196 mRNAs are differentially expressed in Ewing sarcoma, with limited expression in MSCs, while 317 genes were up-regulated in the MSC lines (Figure 1B). This includes known Ewing sarcoma associated proteins such as CD99 (Dataset S2) but CD99 is known as a Ewing sarcoma marker and further CD99 mRNA expression did not fulfill the selectivity criteria used in our study.

The most highly differentially expressed candidate mRNAs in Ewing sarcoma were selected using more stringent criteria, i.e. high expression levels in all three EWS cell lines and no (or very low) expression in both MSC lines, excluding genes with high variability in their expression levels between the Ewing cell lines. This derived a set of ten genes in Ewing sarcoma and four genes in the MSC that fit these criteria (SI Appendix, Table S2); these mRNAs are also indicated in Figure 1B. The attribution of the expression of the 14 genes was confirmed using by real time PCR (qRT-PCR) using cDNA made from the three Ewing cell lines and the two MSC lines. All ten candidate target genes expressed in the Ewing cell lines could not be detected in the MSC lines, while the four MSC-specific cells were detected in the MSC lines, but not in the Ewing cell lines (Figure 2).

LINGO1 is a novel biomarker for Ewing sarcoma tumours

One of the mRNAs observed in the Ewing RNA-seq group encodes the leucine repeat and Ig domain containing protein LINGO1. The protein is expressed in neuronal tissue and is naturally part of the Nogo receptor (13, 18). A most striking characteristic of LINGO1 is its large and well-characterized extracellular domain (19).

This made LINGO1 stand out as a potential new biomarker and drug target in Ewing sarcoma. The expression of *LINGO1* mRNA was studied in a larger panel of Ewing sarcoma cell lines, all of which carry the characteristic chromosomal translocations causing *EWS*- fusion genes. qRT-PCR analysis of eight Ewing sarcoma cell lines and three MSC lines shows that *LINGO1* mRNA could be detected in all of the EWS lines, whereas *LINGO1* was not detected in the MSC lines (Figure 3A).

The analysis of the surfaceome by mRNA expression is a surrogate for the actual proteome and we therefore determined LINGO1 protein levels in the same panel of cell lines using Western blot analysis. The upper panels in Figure 3B shows LINGO1 protein is detected in all of the Ewing cell lines, but not in the MSCs. We verified the presence of the EWS-FLI1 fusion protein by Western blotting with anti-FLI1 antibody (Figure 3B, right) or by Western blotting with anti-EWS in the three Ewing cell lines expressing the alternative fusion proteins EWS-ERG (TTC466 and RM82) or EWS-FEV (STAET10) (Figure 3B, left). In this analysis, the upper band represents cellular EWS protein, while the bottom band represents the fusion protein.

LINGO1 is expressed in all the Ewing sarcoma cell lines tested. The spectrum of primary Ewing sarcoma patient expression was analyzed using tissue microarrays of paraffin-embedded cores taken from tumour biopsies. CD99 is a known immune-histochemical marker of EWS and served as a positive control. LINGO1 expression levels were quantified according to staining intensities. Fifty six patient samples were analyzed in total and we detected LINGO1 expression in 91% of these samples. Twelve samples showed weak, twenty five samples moderate and fourteen samples strong LINGO1 expression levels (SI Appendix, Table S3). Figure 3C shows examples of weak, moderate and strong staining intensities (designated patient 1, 2 and 3 respectively) in different Ewing sarcoma patient samples, all of which are clearly positive for CD99 expression. Paraffin-embedded MSCs from our cultured lines served as negative controls in this experiment and show only background staining (Figure 3C). *LINGO1* RNA expressing was compared in the A673 Ewing cell

line with RNA from a rhabdomyosarcoma cell line and two neuroblastoma lines by q-RT-PCR (SI Appendix, Figure S3). We used the previously identified Ewing sarcoma marker ITM2A (20) (11) and KCNN1 that we have found in our study (SI Appendix, Table S2). *LINGO1* and *ITM2A* expression parallel each other and only show expression A673 and not in the rhabdomyosarcoma or neuroblastomas. KCNN1 on the other hand is relatively highly expressed in the neuroblastoma lines and thus clearly not only in Ewing sarcomas. A micro-array analysis studying expression in several Ewing sarcoma cell lines and other sarcoma types (including RMS) shows LINGO1 is consistently and significantly up-regulated only in Ewing sarcoma (21).

The normal expression of LINGO1 has been reported to be restricted to some neuronal cells and precursors beyond the blood brain barrier (BBB) (13, 22). Further, there are human RNA data showing negligible expression outside the CNS http://www.proteinatlas.org/ENSG00000169783-LINGO1/tissue . We confirmed that in mouse tissues, using RT-PCR, we only observed a LINGO1 product from brain tissue and not from a range of somatic tissues (SI Appendix, Figure S4). The unexpected expression of *LINGO1* in Ewing sarcoma suggests that there may be a relationship to the *EWS* gene fusion that comes from the consistent chromosomal translocation in these tumours. Stable MSC lines were established in which *EWS-FLI1* and *eGFP* expression are driven by a bidirectional doxycycline inducible promoter. Stable clones were induced with 1µg/ml doxycycline and 5 clones displaying high eGFP induction levels were selected. Western analysis with anti-FLI1 antibody to detect the EWS-FLI1 fusion protein and quantitative RT-PCR of *EWS-FLI1* mRNA revealed high induction levels of EWS-FLI1 in the selected clones. (SI Appendix, Figure S5A, B). All five inducible MSC clones showed stimulation of *LINGO1* expression, to a greater or lesser extent compared to the induction of *NR0B1* (the gene encoding DAX1 that is a direct target of EWS-FLI1 (23) expression (SI Appendix, Figure S5C, D). These data endorse *LINGO1* as part of the EWS-FLI1 transcriptome landscape.

LINGO1 is expressed on the surface of Ewing sarcoma cell lines and internalizes when bound by IgG

The immuno-histochemistry of the Ewing sarcoma tumours show surface staining with anti-LINGO1 monoclonal antibody (Figure 3C). Surface expressed LINGO1 protein was confirmed using flow cytometry with eight different Ewing sarcoma cell lines compared to the three MSC lines (Figure 4). All eight Ewing sarcoma cell lines have LINGO1 on their cell surface. These data demonstrate that LINGO1 is a differentially expressed cell surface protein in Ewing sarcoma making it a potentially useful target for antibody-based therapies, such as delivery of Antibody Drug Conjugates (ADC). In order to achieve cell toxicity by ADC methods, the reagent has to bind to a cell surface target and consequently become internalized to deliver the drug. LINGO1 internalization was examined by co-staining the Ewing cell line A673 with FITC-coupled anti-LINGO1 antibody and antibodies recognizing either the early endosome (EEA1) or the lysosome (LAMP1) proteins. Following incubation either at 4°C or at 37°C for 1.5 hours, cells were fixed, permeabilized, and incubated with Alexa Fluor 594 coupled anti-rabbit secondary antibodies, and analyzed by confocal microscopy. We found the anti-LINGO1 antibody was mainly localized at the plasma membrane of cells incubated at 4°C (Figure 5A and B, top) while the antibodies binding EEA1 or LAMP1 were internal to the cells After 1.5 hours incubation at 37°C, the majority of anti-LINGO1 antibody has been internalized and can be seen in the cytoplasm of the cells (Figure 5A and B, bottom). This has been visualized by time-lapse video shown in the SI Appendix, Video 1. A673 cells were incubated (37°C in DMEM medium with 10% FBS) with anti-LINGO1 Alexa Fluor® 488 and green fluorescence time-lapse images overlaid with labelled acidic organelles (lysosomes) using LysoTracker® Deep Red (red). Cells were incubated at and images were collected at 25 seconds intervals for 1 hour using a DeltaVision Elite Imaging System. After internalization, the anti-LINGO1 partially co-localized with the early

endosomal and the lysosomal markers suggesting that LINGO1 protein is internalized via the endosome-lysosome pathway, following binding with the anti-LINGO1 bivalent IgG antibody. As a negative controls, MSCV4 and HEK293 cells (non-expressers of LINGO1 antigen) were incubated with or without anti-LINGO1 Alexa Fluor® 488 (SI Appendix, Videos 2, 3, 4).

LINGO1 can be used as target for Ewing sarcoma cell killing

The unique cell surface expression of LINGO1 protein in Ewing sarcoma compared to other somatic tissues and the phenomenon of antibody-mediated internalization suggests that the protein can serve as a mediator of cell killing using antibody drug conjugates (ADC). This was analyzed using doxorubicin in ADC assays with the A673 Ewing line since these cells are known to be doxorubicin-sensitive (24). A dose response analysis confirmed doxorubicin-sensitivity at about 50uM (Figure 6A). The effect of anti-LINGO1 ADC was assayed in A673 (Figure 6B) by coupling doxorubicin to anti-LINGO1 antibody Li81 and incubating either A673 cells or the MSC line 5Blast for up to 72 hours. A673 cells were also incubated with Li81 antibody alone without conjugated doxorubicin. Cell death was evident in the A673 cells incubated with the anti-LINGO1 ADC at 48 hours resulting in about 25% toxicity within the culture at 72 hours (Figure 6B). The ADC did not affect the percentage viability of the MSC cells, that lack LINGO1 expression, nor did the antibody alone affect the A673 viability.

Discussion

The cancer cell surfaceome

The use of antibody therapies in human diseases, in particular cancer, is gaining importance but a major technical challenge is finding the few cell surface proteins that provide distinguishing marks for specific cancer types. In addition, with the advent of new technologies for tumour targeting, such as invoking T cell responses with chimaeric antigen receptors (25), the requirement for specific tumour cell surface markers has become a critical component for implementing these methods. There are various possible approaches to determining the cell surface proteomes of a tumour types but an under-exploited approach is to utilize the genome sequence and associated annotation resources that provides both the full gamut of human genes, their likely splice variants and their protein products. By analyzing these sequence data, it is possible to classify gene products in terms of likely function but critically also subcellular location, in particular whether the proteins are likely to be at the cell surface.

We have utilized the human genome sequence data and motifs to assort the genes into those encoding proteins that are known to be, or have a likelihood based on the presence of a membrane motif, cell surface proteins (the surfaceome database). This builds on previously published analysis (3) and now includes more comprehensive sets of proteins (all annotated gene products) and respective annotations (including both Uniprot annotations and amino-acid sequence feature predictions). We have assorted the proteins according to likelihood of cell surface expression as gold (known cell surface such as CD markers), silver (multiple independent feature predictions or annotations) and bronze (single feature prediction or annotation). This surfaceome database resource can be used to interrogate next generation deep RNA-seq data from a particular cell type to produce a cell-associated surfaceome. These data can be cell-type specific if suitable control RNA-seq data are available to allow subtraction of expression profiles (subtractive RNA-seq). This approach is

applicable to human disease studies such as cancer biology and auto-immunity but also to developmental biology where cell surface changes influence cell fate (26).

LINGO1 is a potential therapy target in Ewing sarcoma

We have applied this approach to assess surfaceome targets in Ewing sarcoma which is an aggressive bone and soft-tissue sarcoma in adolescents and young adults. Current treatments involve intensive chemotherapy, radiotherapy and radical surgery. In an attempt to invoke molecular biology methods to identify new approaches to Ewing sarcoma therapy, we have generated RNA-seq data from tumour cells and compared these with data from mesenchymal stem cell lines to carry out subtractive RNA-seq. Among our list of differentially expressed candidate surfaceome targets were members of immunoglobulin-like domains, the G protein-coupled receptor superfamily, ion channels as well as ion transporters (SI Appendix, Table S2) and includes seven proteins, that have not previously been highlighted as EWS surface targets (LINGO1; KCNN1, CDH23, ADRA1D, SLC24A3, CACNA1H, SLC29A4). A recent review has summarized available Ewing sarcoma transcriptome data (27). We have used our surfaceome database mining strategy to assess the published RNA-seq datasets (28, 29) (this excludes one dataset that is currently not publically available due to patient confidentiality (30)) and we found excellent concordance with the expression values found for our Ewing sarcoma surfaceome candidates shown in SI Appendix, Table S2. In both datasets, the expression level of the Ewing's up-regulated candidates positively correlates with the expression level of the translocation fusion *EWS-FLI1*, and conversely mRNAs up-regulated in MSC negatively correlates with *EWS-FLI1*. Amongst these outstanding proteins is LINGO1 that consists of 620 amino acids with a large extracellular domain, the structure of which has been elucidated by crystallography (19), and which is displayed at the cell surface, thereby favouring immune regulation approaches mediated by antibody binding.

LINGO1 is a prominent molecule since an extracellular immunoglobulin domain and a leucine repeat domain have provided antigenic epitopes for antibody derivation, including an anti-LINGO1 antibody that is currently in clinical trials (31). LINGO1 is a component of the Nogo receptor signaling complex and plays a role in regulation of neuronal survival, axon regeneration, oligodendrocyte differentiation and myelination (13, 18). LINGO1 is exclusively expressed in the CNS (32) as we have verified with RT-PCR analysis of RNA from various mouse tissues (SI Appendix, figure S4). Brain-expressed LINGO-1 is largely protected from circulating antibodies by the blood brain barrier. The restricted expression pattern and accessible extracellular domain make LINGO1 a attractive target for antibody-based therapies.

Our results show that an antibody binding to the extracellular domain of LINGO1 can induce cell killing when the anti-LINGO1 antibody carries a cytotoxic drug as an ADC. These immunological properties are presumably due to the efficient internalization of the cell surface LINGO1 protein on binding to the bivalent IgG1 antibody used. Thus ADC is a most promising method for treatment of Ewing sarcoma. In Ewing, the site and size of the primary tumour are prognostic factors for outcome but the most important factor is the presence of metastatic disease at presentation which is an adverse factor. About a quarter of patients present with disseminated disease and the front line therapy (surgery and chemotherapy) results in low 5 year survival. Recurrent disease is associated with a very poor prognosis and new therapeutic approaches are required (33). The use of targeting strategies to LINGO1 is a novel potential approach to improve this outcome.

The expression of LINGO1 on brain cells is an issue with the use of antibodies carrying toxic drugs. The blood-brain barrier is an effective gross prevention of blood supply because of the tight junctions of the vessel endothelium but some macromolecule movement is possible. Thus ADC with anti-LINGO1 would need to be tested to assess this possible problem. Anti-LINGO1 bi-specific

antibodies could be an important approach to avoid toxicity issues. Special formulations of immuno-nanoparticles with drug payloads are developmental options that can also potentially overcome any toxicity issues.

Materials and Methods

RNA-seq library preparation and data acquisition

For RNA-seq analysis of the MSC lines 5H and 4+v and the EWS cell lines, A673, TC-32 and TTC-466, total RNA was extracted and sequencing libraries prepared before single end deep sequencing using an Illumina GA IIx to obtain 80bp reads. RNA-seq data and surfaceome database analysis was carried out as described in the Supporting Appendix. Other general methods and associated references are available online in the the Supporting Appendix.

Further supporting information and data files are available online.

Acknowledgements

This work was supported by grants from the Medical Research Council (MR/J000612/1), Wellcome Trust (099246/Z/12/Z) and Leukaemia and Lymphoma Research (12051). We wish to thank Dr. Juan Funes and Prof. Chris Boshoff for the MSC cell lines and Prof. Bass Hassan and Dr. Harriet Brandford-White for the Ewing sarcoma cell lines CHP100, RM82, STAET1, STAET10 and SKNMC cells and the Ewing sarcoma tissue micro-arrays. We would also like to thanks Prof. Andrew Bradbury for the IgG expression vectors.

Conflict of interest statement

The authors declare that they have no conflict of interest.

Author's contributions

Originator of project: THR

Participated in research design: JT, HP, SH, LFS, JZ, THR

Conducted experiments: JT, SH, CB, WB, JZ

Contributed new reagents or analytic tools: HP, LFS

Performed data analysis: JT, HP, SH, LFS, JZ, THR

Wrote or contributed to the writing of the manuscript: all authors

References

1. Weiner LM, Surana R, & Wang S (2010) Monoclonal antibodies: versatile platforms for cancer immunotherapy. *Nature reviews. Immunology* 10(5):317-327.
2. Wang Z, Gerstein M, & Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews. Genetics* 10(1):57-63.
3. da Cunha JP*, et al.* (2009) Bioinformatics construction of the human cell surfaceome. *Proceedings of the National Academy of Sciences of the United States of America* 106(39):16752-16757.
4. Sankar S & Lessnick SL (2011) Promiscuous partnerships in Ewing's sarcoma. *Cancer genetics* 204(7):351-365.
5. Terrier P, Llombart-Bosch A, & Contesso G (1996) Small round blue cell tumors in bone: prognostic factors correlated to Ewing's sarcoma and neuroectodermal tumors. *Seminars in diagnostic pathology* 13(3):250-257.
6. Abed R & Grimer R (2010) Surgical modalities in the treatment of bone sarcoma in children. *Cancer treatment reviews* 36(4):342-347.
7. Potratz J, Dirksen U, Jurgens H, & Craft A (2012) Ewing sarcoma: clinical state-of-the-art. *Pediatric hematology and oncology* 29(1):1-11.
8. Gorlick R, Janeway K, Lessnick S, Randall RL, & Marina N (2013) Children's Oncology Group's 2013 blueprint for research: bone tumors. *Pediatric blood & cancer* 60(6):1009-1015.
9. Deneen B & Denny CT (2001) Loss of p16 pathways stabilizes EWS/FLI1 expression and complements EWS/FLI1 mediated transformation. *Oncogene* 20(46):6731-6741.
10. Lessnick SL, Dacwag CS, & Golub TR (2002) The Ewing's sarcoma oncoprotein EWS/FLI induces a p53-dependent growth arrest in primary human fibroblasts. *Cancer cell* 1(4):393-401.

11. Riggi N*, et al.* (2008) EWS-FLI-1 expression triggers a Ewing's sarcoma initiation program in primary human mesenchymal stem cells. *Cancer research* 68(7):2176-2185.

12. Tirode F*, et al.* (2007) Mesenchymal stem cell features of Ewing tumors. *Cancer cell* 11(5):421-429.

13. Mi S*, et al.* (2004) LINGO-1 is a component of the Nogo-66 receptor/p75 signaling complex. *Nature neuroscience* 7(3):221-228.

14. Nagalakshmi U*, et al.* (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* 320(5881):1344-1349.

15. Wilhelm BT*, et al.* (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* 453(7199):1239-1243.

16. Mortazavi A, Williams BA, McCue K, Schaeffer L, & Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods* 5(7):621-628.

17. Wild L, Funes JM, Boshoff C, & Flanagan JM (2010) In vitro transformation of mesenchymal stem cells induces gradual genomic hypomethylation. *Carcinogenesis* 31(10):1854-1862.

18. Mi S*, et al.* (2005) LINGO-1 negatively regulates myelination by oligodendrocytes. *Nature neuroscience* 8(6):745-751.

19. Mosyak L*, et al.* (2006) The structure of the Lingo-1 ectodomain, a module implicated in central nervous system repair inhibition. *The Journal of biological chemistry* 281(47):36378-36390.

20. Staege MS*, et al.* (2004) DNA microarrays reveal relationship of Ewing family tumors to both endothelial and fetal neural crest-derived cells and define novel targets. *Cancer Res* 64(22):8213-8221.

21. Teicher BA*, et al.* (2015) Sarcoma Cell Line Screen of Oncology Drugs and Investigational Agents Identifies Patterns Associated with Gene and microRNA Expression. *Mol Cancer Ther* 14(11):2452-2462.

22. Carim-Todd L, Escarceller M, Estivill X, & Sumoy L (2003) LRRN6A/LERN1 (leucine-rich repeat neuronal protein 1), a novel gene with enriched expression in limbic system and neocortex. *The European journal of neuroscience* 18(12):3167-3182.

23. Garcia-Aragoncillo E*, et al.* (2008) DAX1, a direct target of EWS/FLI1 oncoprotein, is a principal regulator of cell-cycle progression in Ewing's tumor cells. *Oncogene* 27(46):6034-6043.

24. Martins AS*, et al.* (2006) Insulin-like growth factor I receptor pathway inhibition by ADW742, alone or in combination with imatinib, doxorubicin, or vincristine, is a novel therapeutic approach in Ewing tumor. *Clin Cancer Res* 12(11 Pt 1):3532-3540.

25. Gill S & June CH (2015) Going viral: chimeric antigen receptor T-cell therapy for hematological malignancies. *Immunological reviews* 263(1):68-89.

26. Rugg-Gunn PJ*, et al.* (2012) Cell-surface proteomics identifies lineage-specific markers of embryo-derived stem cells. *Developmental cell* 22(4):887-901.

27. Sand LG, Szuhai K, & Hogendoorn PC (2015) Sequencing Overview of Ewing Sarcoma: A Journey across Genomic, Epigenomic and Transcriptomic Landscapes. *Int J Mol Sci* 16(7):16176-16215.

28. Marques Howarth M*, et al.* (2014) Long noncoding RNA EWSAT1-mediated gene repression facilitates Ewing sarcoma oncogenesis. *J Clin Invest* 124(12):5275-5290.

29. Riggi N*, et al.* (2014) EWS-FLI1 utilizes divergent chromatin remodeling mechanisms to directly activate or repress enhancer elements in Ewing sarcoma. *Cancer Cell* 26(5):668-681.

30. Brohl AS*, et al.* (2014) The genomic landscape of the Ewing Sarcoma family of tumors reveals recurrent STAG2 mutation. *PLoS Genet* 10(7):e1004475.

31. Tran JQ*, et al.* (2014) Randomized phase I trials of the safety/tolerability of anti-LINGO-1 monoclonal antibody BIIB033. *Neurology(R) neuroimmunology & neuroinflammation* 1(2):e18.

32. Carim-Todd L, Escarceller M, Estivill X, & Sumoy L (2003) LRRN6A/LERN1 (leucine-rich repeat neuronal protein 1), a novel gene with enriched expression in limbic system and neocortex. *The European journal of neuroscience* 18(12):3167-3182.

33. Gaspar N*, et al.* (2015) Ewing Sarcoma: Current Management and Future Approaches Through Collaboration. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*.

34. Schumacher FF*, et al.* (2014) Next generation maleimides enable the controlled assembly of antibody-drug conjugates via native disulfide bond bridging. *Organic & biomolecular chemistry* 12(37):7261-7269.

Figure legends

Figure 1. Gene expression quantification by RNA-seq reveals differentially expressed genes in Ewing sarcoma

Mean expression values in the MSC lines are shown on the x axis and in the Ewing sarcoma cell lines on the y axis. Values are transcripts per million (TPM). Plots are shown for the entire transcriptome (A) or mRNA transcripts encoding surfaceome proteins (B). Each dot represents an mRNA transcript. Transcripts found to be significantly up-regulated in Ewing sarcoma or MSC are shown with red or blue dots respectively (for each transcript we carry out a modified t-test using the three EWS values and the two MSC values. From this test we obtain a p-value, corrected for multiple testing. A transcript is marked as up-regulated or down-regulated only if the corrected p-value is smaller than 0.05). The numbers, in panel B, indicate the surfaceome candidate target genes identified in Ewing sarcoma (1 -10) and MSC-specific genes (11-14) (see SI Appendix, Table S2).

Figure 2. Analysis of candidate cell surface gene expression levels by qRT-PCR

RNA isolated from the Ewing sarcoma cell lines A673, TC-32 and TTC466 and two MSC lines (5H and 4+v) was reverse transcribed into cDNA and analysed by qRT-PCR. Expression levels were normalized against the house-keeping gene *GAPDH* and the Ewing sarcoma cell line A673 was used as a reference. Relative expression levels are given as $2^{-\Delta\Delta Ct}$ with $\Delta Ct = Ct_{LINGO1} - Ct_{GAPDH}$, $\Delta\Delta Ct = \Delta Ct_{\ Sample} - \Delta Ct_{A673}$. The error bars represent the 95% confidence interval of the RQ (relative quantity) value.

Figure 3. LINGO1 is differentially expressed in Ewing Sarcoma

The association of LINGO1 mRNA and protein in Ewing sarcoma cells was confirmed by qRT-PCR and immunoblotting.

A. qRT-PCR analysis of eight Ewing sarcoma cell lines and three MSC lines. *LINGO1* expression levels were normalized against the house-keeping gene *GAPDH* and the Ewing sarcoma cell line A673 was used as a reference. These are set at 1, i.e. A673 =100%. Relative expression levels are given as RQ (relative quantification) = $2^{-\Delta\Delta Ct}$ with $\Delta Ct = Ct_{LINGO1} - Ct_{GAPDH}$, $\Delta\Delta Ct = \Delta Ct_{Sample} - \Delta Ct_{A673}$. The error bars represent the 95% confidence interval of the RQ value.

B. Western blot analysis of LINGO1 protein. Lysates of Ewing sarcoma and MSC cells were fractionated by SDS-PAGE electrophoresis and analysed by Western blotting using the anti-LINGO1 antibody (Abcam). Expression of EWS-FLI1 was shown in A673, TC-32, CHP100, SKNMC, STAET1 (right hand panel) using an anti-FLI1 antibody. Expression of either the alternative fusion protein EWS-ERG in TTC-466, RM82 or the EWS-FEV fusion in STAET100 (left hand panel) was demonstrated using an anti-ERG antibody. Actin served as a protein loading control.

C. LINGO1 is expressed in primary Ewing Sarcoma patient samples. Tissue microarrays containing cores from tumour biopsies were analyzed by immunohistochemistry using a mixture of two AlexaFluor 488-coupled (green) anti-LINGO1 antibodies (Abcam and Millipore) and an Alexa-Fluor594-coupled (red) anti-CD99 antibody (Thermo Scientific). Images were acquired by confocal laser scanning microscopy and analyzed using ImageJ software. LINGO1 expression levels were estimated based on staining intensities. The figure shows examples of weak (patient 778), medium (patient 786) and strong (patient 811) LINGO1 staining intensities (summarized data are given in SI Appendix, table S3). Paraffin embedded MSC tissue culture cells served as a negative control. Scale bars = 20µm.

Figure 4: LINGO1 is expressed on the surface of Ewing Sarcoma cell lines

LINGO1 surface expression in (A) EWS and (B) MSC cell lines was analyzed by FACS using the anti-LINGO1 antibody Li81 and a FITC-conjugated anti-human IgG secondary antibody. The x-axis shows $\log_{10}$ fluorescence intensities for LINGO1

antibody binding (blue) while the y-axis shows cell counts normalized to maximum of cells collected for each cell line. The cell staining with second antibody-only is shown in red.

Figure 5: LINGO1 protein internalizes on Ewing Sarcoma cells when bound by bivalent antibody

LINGO1 is internalized and localizes to the early endosomes and lysosomes after binding anti-LINGO1 antibody. A673 Ewing sarcoma cells was treated with Alexa 488-labeled (green) anti-LINGO1 Li81 antibody and incubated for 1.5 hours at 4^0C or 37^0C. Cells were fixed and incubated with monoclonal rabbit antibody binding either the endosomal marker EEA1 (panel A) or the lysosomal marker LAMP1 (panel B) followed by incubation with Alexa 594-labeled (red) secondary anti-rabbit antibodies. Cell nuclei were stained using DAPI (blue).

Figure 6. Effect of doxorubicin anti-LINGO1 antibody drug conjugate on viability of Ewing sarcoma cells

A. The Ewing sarcoma cell line A673 expressing LINGO1 surface antigen was tested for sensitivity to doxorubicin. Cells were incubated in the absence or presence of increasing concentrations of doxorubicin and the cell viability was counted using the PrestoBlue assay at 48 hours.

B. Li81 anti-LINGO1 antibody was converted to an ADC using next generation maleimide (34) and using in cell killing assays. A673 and MSC control cells were incubated for 48 or 72 hours in the presence of 100ug/ml Li81-ADC or Li81 alone. After the indicated incubation time, the cell viability was assayed using the PrestoBlue method. Values are normalized to cell number and viability of untreated cells.