Experimental design

The following step-by-step procedure covers the immunostaining of formaldehyde-fixed, paraffin-embedded human teratoma tissue for the detection of mouse cyclophilin A to assess host contribution in hESC-derived xenografts. It essentially follows standard methodology for IHC and immunofluorescence which we carefully adjusted for combination with the candidate cyclophilin A antibody. The resultant fluorescent IHC protocol is straightforward, technically feasible and generally well doable in any laboratory equipped for basic molecular or cell biological techniques. Additional considerations regarding the overall experimental set-up, including (i) preliminary antibody validation, (ii) teratoma formation and tissue sampling, and (iii) the immunostaining protocol itself are outlined in the subsequent paragraphs and briefly discussed.

Preliminary antibody validation. The identified cyclophilin A (D2Y4M) rabbit monoclonal antibody is the key element of the procedure, and its preliminary validation for this particular application represents one of the most crucial steps within this protocol, as it does for any antibody-based approach\textsuperscript{52,53}. The supplier-provided validation of the antibody basically covers its target specificity in immunoblotting applications, and its reactivity towards the mouse protein, which is used for immunization. However, its performance in IHC or related techniques such as immunofluorescence on cultured cells or flow cytometry, as well as its actual cross-reactivity towards the human ortholog remained entirely elusive at this point. Preliminary antibody characterization with particular regard to these two criteria was thus the rate-limiting step in the endeavor to implement this protocol.

As described above, prior to testing on tissue, antibody performance was thoroughly evaluated at the cellular level upon different applications involving the denatured and native protein in its intracellular and secreted form (Figs 2 and 3; Supplementary Figs 1-3). The tested cell panel includes cells of mouse and human origin, with the majority being common in labstocks or commercially available. It was assembled to cover a small range of distinct cell types to roughly account for the different cellular origins in the experiment. That is the tumorigenic cells themselves (hPSC or somatic tumor cells), the differentiated cells of the teratoma and the host cells of the teratoma/tumor microenvironment (e.g. fibroblasts and macrophages) (Table 1). Next to monoculture, human and mouse cells were tested upon in vitro co-culture at defined ratios to estimate antibody performance in a mixed-species background (Fig. 3b,c; Supplementary Fig. 3). In summary, these experiments provide valuable information on antibody characteristics such as sensitivity and the ability to recognize the target epitope upon distinct conformational or preservative states including formaldehyde fixation. But most importantly, they confirm the hypothesized species specificity of the cyclophilin A antibody (Fig. 1b), that is being "mouse-but-not-human"-reactive. At this point, it should be noted that the results of our preliminary validation experiments regarding species cross-reactivity are in line with the later obtained internal testing data provided by the supplier. Information on reactivity towards other species (e.g. rat) might be available upon request (support@cellsignal.com). Additional validation experiments including the comparison with other houskeeping genes (Fig. 2b) as well as complementary analyses...
using human-specific antibodies (Fig. 2b; Fig. 3b,c; Table 2) further substantiate the suitability of the antibody respectively approach for the intended application.

Immunoblotting was the method of choice to verify species specificity in a sensitive manner across multiple cell lines under identical conditions. Basically, the detection of a specific band in immunoblotting does neither guarantee nor indicate that the antibody is going to perform equally well in IHC. Sample preparation for reducing/denaturing electrophoresis is harsh and might expose antigens which are otherwise inaccessible if proteins remain in a more native, three-dimensional state\textsuperscript{54,55}. However, the presence or absence of additional bands is usually a good indicator to rate antibody performance in terms of cross-reactivity towards unrelated epitopes, which contributes to non-specific staining not readily visible in in situ approaches. In this respect, the cyclophilin A (D2Y4M) rabbit monoclonal antibody gives optimum results, with no apparent signs of cross-reactivity towards other proteins in both, human and mouse (Fig. 2a; Supplementary Fig. 1a; Supplementary Fig. 2a). Finally, the fact that the antibody recognizes the epitope upon formaldehyde fixation (Fig. 3) does not only anticipate its basic suitability for IHC on paraffin-embedded tissue, but provides a reference pattern of subcellular localization to which IHC results can be compared. Note that the obtained cytoplasmic-nuclear localization pattern is consistent with the known subcellular expression profile of mouse cyclophilin A (http://www.informatics.jax.org/marker/key/12618).

This antibody validation profile is a self-contained building block of the procedure and represents a preliminary step which, in principle, does not require rerun or any kind of adjustments in order to complete the protocol. However, some users might want to prevalidate the antibody in a context-specific manner for use in their cellular system under study. To do so, all details on the used cell lines and reagents to serve as positive and negative controls, are summarized in Tables 1-3.

**Teratoma formation and tissue sampling.** The protocol has been established upon injection of 1x10^6 hESC into the gastrocnemius muscle of immunodeficient NOD-scid IL2R\textsuperscript{null} mice (Boxes 1 and 2)\textsuperscript{7,56} to produce early-stage (4 weeks post-injection) and mature (8 weeks post-injection) teratomas for subsequent studies of the teratoma microenvironment. Over the course of experiments, we found the growth stage of hESC at the time of harvest to be critical to obtain consistent results. Dense cultures tend to slightly reduce the incidence of teratoma formation and increase the time required to develop a tumor. Other key parameters including mouse age, sex and the number of injected cells are kept constant throughout to ensure reproducibility in and between experiments. This is especially important if experimental variables are introduced to study their effect on host cell contribution\textsuperscript{7}. Due to the high specificity and sensitivity of the cyclophilin A antibody in the tested application and the chosen detection system, a very robust signal and an extremely low, virtually zero, failure rate has to be assumed. Hence, technical replicates, e.g. the dual staining of consecutive sections, can be entirely omitted in favor of biological replicates. To our experience, usually six to nine teratomas from three independent injection experiments are sufficient to yield a representative picture of the associated microenvironment and to establish a baseline for the tumorigenic cell line of interest.
The mouse hind limb respectively gastrocnemius muscle is a site well suited for teratoma growth and experimentation. It is readily accessible for cell engraftment and tumor explantation, with the former being entirely free from surgical procedures, and the latter requiring only simple surgical techniques. Moreover, tumor growth can be easily monitored by mere palpation. Next to these practical considerations, it turned out that this transplantation site is particularly suitable to not only assess host cell infiltration accumulating at the teratoma boundary and in the teratoma itself, but to actually trace it over the teratoma-adjacent tissue. The skeletal muscle tissue provides a solid cellular matrix in which the cystic teratoma is firmly embedded, and which is visibly remodeled upon host cell recruitment. In the course of teratoma explantation, it can be more or less freely excised to study mouse cell infiltration over a wider distance. Beyond that, skeletal muscle tissue is highly vascularized and shows a unique cellular structure. Large, tightly packed muscle cells are sparsely interspersed with interstitial cells, mainly fibroblasts, lymphatic and blood vessels, nerves, and mesenchymal progenitors. Due to these particular morphological properties, the engrafted human tissue and the invaded host cells stand out distinctively against the surrounding tissue, the latter being readily identified via mouse-specific staining of cyclophilin A (Fig. 5; Supplementary Figs. 4 and 5).

The teratoma formation protocol uses Matrigel as a solidifying support matrix to promote cell engraftment in the hind limb post-injection (Boxes 1 and 2). It counteracts the random dissemination of transplanted cells, holding them together at the transplantation site which is crucial for teratoma initiation. The effect of Matrigel can be enhanced by anesthetizing the mice for injection and keeping them on a heating pad to allow the cell/Matrigel mixture to solidify completely, forming a compact injection plug. However, using the simple approach outlined in Box 2, we consistently achieved a teratoma formation efficiency of 80-100%.

Standard Matrigel contains significant amounts of growth factors, with some of them being well characterized for their invasion-promoting effects. To minimize the impact of non stem cell-derived factors on teratoma formation and especially host cell activation, growth factor-reduced Matrigel is used. It holds only one third to one half of the growth factor content of standard formulations, while retaining the original composition of extracellular matrix components. To be able to (i) distinguish stem cell-induced effects from unrelated events, that is those associated with the application of Matrigel and the injection procedure itself, and to (ii) rule out the unlikely event of tumor formation by Matrigel, a randomized control group (≥2 mice per injection experiment) is grafted with Matrigel only (Box 2). This control serves as a valuable reference for what normal, unaffected tissue looks like and provides the baseline to which results are compared. This could be especially relevant if very early events post-injection are studied which, at least to some extent, might reflect the injection procedure itself. In summary, we found the Matrigel-only control to be essential to accurately and sensitively evaluate the impact of stem cell application on the host. This is nicely shown by the identification of less pronounced effects in the teratoma-adjacent tissue, e.g. the subtle accumulation of activated host cells at more distant sites which presumably precedes their recruitment into the teratoma.

The simple approach described above is part of a tissue sampling strategy on a larger scale, which we specifically developed to study the teratoma microenvironment at both, the molecular and
functional level. For this purpose, collected tissues are divided into portions and either formaldehyde-fixed, snap-frozen in liquid nitrogen or left untreated for subsequent use in IHC, immunoblotting and tissue culture, respectively. For details on the procedure and the type of collected tissues, see reference 7. Similar to the validation at the cellular level, immunoblotting of collected tissues is a feasible approach to pre-assess cyclophilin A antibody reactivity towards unrelated proteins of different molecular weight. This is particularly useful in the context of hESC-derived xenografts, where the grafted cells undergo massive differentiation to generate derivatives of all three embryonic germ layers and thus do no longer reflect the initially transplanted and characterized cells. Representative immunoblotting data in Figure 4b demonstrate the absence of additional, unspecific bands in both, the skeletal muscle tissue at the transplantation site and the host-infiltrated teratoma tissue. These results are in line with the preliminary antibody validation data described above, and largely exclude non-specific binding at the tissue level. Details on the preparation of tissue lysates for subsequent SDS-Page and immunoblotting are summarized in Box 3.

**Tissue processing and immunostaining.** After explantation and fixation in formaldehyde, tissues are further processed for embedding in paraffin and sectioning. To do so, samples are progressively dehydrated in increasing concentrations of ethanol to remove the water and the fixative, before being cleared in xylene which is finally replaced by molten paraffin wax. Permeated tissues are then embedded in extra paraffin forming the later "tissue block" which provides a robust support matrix during sectioning. Finally, the tissue is sectioned using a microtome, and mounted onto positively charged adhesion slides. As for this protocol, the dehydration, clearing and embedding steps are usually carried out following broadly available standard procedures, mostly by making use of fully or semi-automated tissue processors and embedding centers which save time and yield consistent results. To our experience, the application of a particular protocol at this stage is not mandatory since slight variations in incubation time, the number of implemented incubation steps or the grading of the ethanol series are well compatible with our protocol and do not significantly affect the outcome of the overall procedure.

Fixation and paraffin-embedding both play crucial roles in preserving tissue morphology and retaining the antigenicity of the target epitope. Aldehydes immobilize and stabilize soluble proteins through chemical crosslinks while maintaining the cellular and subcellular structure of the sample. This process does not only impact the basic antigenicity of the epitope by introducing changes in the chemical properties and the three-dimensional conformation of the proteins, but is also known to restrict antibody access to the antigen, a phenomenon called antigen or epitope masking\textsuperscript{54,55}. Antigen retrieval methods aim at reversing these crosslinks in order to re-expose the antigen for the antibody to bind and to counteract the fixation effects on the target epitope. The occurrence of antigen masking and the optimum technique for retrieval depend on several factors including the tissue, the fixation protocol, the antigen target and/or the antibody. In summary, it is important to bring to mind that the requirements for antigen retrieval are not specific to the protein itself but arise from the unique interaction of the epitope and the corresponding antibody, and hence have to be evaluated and optimized for every new antibody.
Freezing is an alternative and commonly used method to preserve tissue for subsequent immunostaining. The staining of frozen sections typically employs shorter protocols requiring less processing steps since there is no need for tissue de- and rehydration or antigen retrieval. Cryosections however are less robust, often thicker than paraffin sections and are generally less efficient in preserving the original tissue structure with high fidelity, the latter being vital for in-depth studies of the teratoma microenvironment. In our hands, the cyclophilin A staining of frozen teratoma sections was associated with a selective and progressive loss of tissue during the staining procedure, making it impossible to obtain a complete picture of host cell recruitment which renders the approach more or less unreliable. This outcome might be due to the distinct textures of the different tissues found in a teratoma which are preserved at varying efficiency and which require protocol optimization by an experienced user with expert knowledge on cryopreservation and -sectioning. Beyond that, paraffin-embedding followed by sectioning and hematoxylin and eosin (H&E) staining is the standard method to verify the formation of mature teratomas containing derivatives of all three embryonic germ layers. This is a prerequisite to establish the pluripotency of a hPSC line in vivo and is crucial to data interpretation. Hence, next to giving superior results in tissue preservation, formaldehyde fixation and paraffin embedding offers the possibility to stain serial sections for H&E and mouse cyclophilin A which allows to roughly assign the pattern of infiltrated host cells to mesodermal, endodermal and ectodermal tissues. Depending on the specific question, this is a convenient and cost-saving approach before proceeding to more costly marker stainings.

Although routinely used in basic research and clinical settings, stand-alone IHC is considered a demanding and challenging technique. This view is based on the fact that the likelihood of misleading artefacts to occur is relatively high if the used antibody did not pass through thorough, application-specific validation. However, once successfully established for a protein of interest, it is a very robust method, not prone to variability in and between experiments. Key parameters, which have to be evaluated and optimized for every antibody, include (i) the antigen retrieval method, (ii) the tissue blocking and permeabilization steps, (iii) the antibody concentration and diluent and (iv) the detection system used to visualize antibody binding. All practical considerations regarding these variables are annotated in the step-by-step procedure and, if vital for the outcome, highlighted as "CRITICAL STEPS".

In brief, for the antibody to be able to bind to its target epitope, the paraffin wax from the embedding and sectioning steps must be completely removed and the tissue has to be rehydrated. The deparaffinization and rehydration procedure used in this protocol is detailed in the step-by-step procedure and is recommended to be implemented without modifications to achieve optimum results. To restore target antigenicity and to allow antibody binding to occur, masked epitopes are usually retrieved through either enzymatic digestion or boiling in an appropriate buffer at defined pH, the latter referred to as "HIER" or heat-induced epitope retrieval. Available HIER protocols mainly differ in the type of buffer and/or boiling device used. The retrieval protocol for the cyclophilin A antibody has been established upon experimental testing, and represents a gentle approach which uses microwave boiling in sodium citrate buffer at low power settings and defined short-term intervals. Importantly, this
protocol performs equally well for other targets respectively antibodies, either in singleplex or cyclophilin A double staining approaches, thereby significantly broadening the range of possible applications. The dual staining of sections for cyclophilin A and a protein of interest does not only allow to identify or define the specific cell type in more detail, but offers the possibility to phenotypically characterize the infiltrating host cells for other features, e.g. their activation status. As an example, the step-by-step procedure describes the optional double staining of teratoma tissue for cyclophilin A and α-smooth muscle actin (SMA) to identify vessel-like structures of host origin.

Next to efficient epitope retrieval, protein blocking to prevent non-specific antibody binding is important to lower background staining and to obtain a good signal-to-noise ratio. Non-specific binding applies to both, primary and secondary antibodies, and is more or less defined as the binding of an antibody to a site other than the intended. This occurs through simple adsorption or due to hydrophobic, charged-based or other type of interactions. A frequent source of non-specific binding in tissues are Fc receptors on immune cells such as monocytes and macrophages, predominantly in organs such as the spleen, the thymus or the blood, but also in the tumor microenvironment. The most commonly used blocking agents are serum, which contains antibodies that readily bind to reactive sites, and bovine serum albumin (BSA). For indirect labeling approaches as used in this protocol, the former has to be matched to the species of the secondary antibody. In our hands, both blocking options give very similar results, making them virtually interchangeable in the procedure. To guarantee antibody access to intracellular including nuclear epitopes, the blocking solution is supplemented with Triton X-100, a relatively harsh detergent which permeabilizes the cell and nuclear membranes to create pores of sufficient size while keeping the cellular and subcellular structure intact.

Key variables that determine primary antibody performance, including antibody concentration, diluent, incubation time and temperature, have been optimized to yield a robust and specific signal at maximum sensitivity and minimum background staining. As mentioned above, indirect detection of the primary antibody as used in our protocol, has some intrinsic limitations for conventional double and triple staining, and also requires additional work steps and controls. However, signal amplification through two or more labeled secondary antibodies binding to each primary antibody also provides increased sensitivity. This allows to compensate for minor differences in expression levels to visualize as many cell types as possible. To reduce non-specific binding, our protocol employs fluorophore-conjugated F(ab')₂ fragment secondary antibodies which lack the Fc portion and are pre-adsorbed against human and mouse serum. Due to their smaller size, they efficiently penetrate the tissue and contribute to sensitivity at an additional level. The fact that the cyclophilin A antibody is raised in rabbit and not in mouse, eliminates the need for an additional blocking step to prevent binding of the secondary antibody to endogenous mouse immunoglobulins in the tissue. However, this has to be considered for the "mouse-on-mouse" detection in cyclophilin A double stainings, especially if high background staining is observed.

In our approach, the inclusion of appropriate controls to distinguish a true staining pattern from an artefact is largely covered by the preliminary validation experiments (Figs 2-4; Supplementary Figs
1-3), which together form the reliable basis to rate the validity of the obtained IHC results. In summary, these key observations include (i) the "mouse-but-not-human" reactivity, (ii) the lack of antibody cross-reactivity towards unrelated proteins, either denatured or native, in both species, (iii) the comparable expression level between distinct cell types and (iv) the subcellular localization pattern.

Classical IHC employs antigen controls to prove target specificity, reagent controls to exclude artefacts caused by the detection system itself, and occasionally tissue controls to evaluate the impact of autofluorescence. The side-by-side staining of sections from formaldehyde-fixed and paraffin-embedded tissues or cells known to either express or lack the protein of interest is a common test to identify false negative and especially false positive staining. The unique nature of our approach however, based on the selectivity of the antibody and the mixed-species character of the sample, allows to confirm antibody specificity right in the tissue under investigation. In this context, the most conclusive demonstration of specificity is the lack of cyclophilin A staining in cells of human origin. This has been verified by counterstaining with a human-specific antibody targeting ubiquitously expressed topoisomerase IIβ\(^7\). These data, similar to those presented in Figs 3b and c, did not show any double-positive cells which largely excludes false positive results. The fact that double-negative cells could not be detected neither, provides some evidence that the cyclophilin A antibody indeed detects the vast majority of contributing host cells, at least in the teratoma itself. Moreover, the obtained localization pattern and the similar expression level between individual cells are both consistent with the \textit{in vitro} validation data and the known expression of the protein, which further substantiates the validity of the staining. Incubation with the primary antibody that has been pre-adsorbed against the peptide immunogen would be another valuable approach to prove target specificity right in the tissue under investigation, but is currently not on-hand due to the lack of a commercially available blocking peptide.

Controls, which are included on a regular basis, are those that monitor the detection system respectively the staining procedure itself. Specifically, the protocol uses non-immune, species- and concentration-matched control immunoglobulins, so called isotype controls, which replace the primary antibody in the procedure. This control is supposed to detect background staining caused by the non-specific interaction of the primary antibody with the tissue due to Fc receptor binding or other types of interactions. Isotype controls are preferentially analysed on serial sections of the same tissue to sustain identical conditions. For the detection procedure to be valid, signals obtained with this control must be clearly distinct from the presumably specific signal, or ideally almost undetectable, as seen with this protocol. If high background staining is observed, a "no-primary-antibody"-control should be performed before proceeding to more extensive troubleshooting. This allows to check whether the secondary is the main source of non-specific staining and is simply done by replacing the primary antibody with antibody diluent. This control is especially reasonable if a secondary other than the recommended is used. If other tissues than mouse skeletal muscle are injected, autofluorescence might contribute to background staining which can be roughly estimated by co-analysing an unstained sample.
Finally, it is essential to note that the potential problems listed in the troubleshooting table are extremely unlikely to occur if the individual steps of the protocol are followed as outlined, and random modifications or shortcuts of the procedure are avoided. Nonetheless, they are summarized to highlight the most critical steps, and to provide a troubleshooting basis for those who lack expert knowledge and practical experience in IHC or related techniques.