Stem Cell Reports
Article

OPEN ACCESS

Preclinical Efficacy Failure of Human CNS-Derived Stem Cells for Use in the Pathway Study of Cervical Spinal Cord Injury

Aileen J. Anderson,[1,2,3,4,*] Katja M. Piltti,[1,3] Mitra J. Hooshmand,[1,3] Rebecca A. Nishi,[1,3] and Brian J. Cummings[1,2,3,4]

[1]Sue & Bill Gross Stem Cell Center
[2]Physical & Medical Rehabilitation
[3]Institute for Memory Impairments & Neurological Disorders
[4]Anatomy & Neurobiology
University of California-Irvine, Irvine, CA 92697, USA
*Correspondence: aja@uci.edu
http://dx.doi.org/10.1016/j.stemcr.2016.12.018

SUMMARY

We previously showed the efficacy of multiple research cell lines (RCLs) of human CNS neural stem cells (HuCNS-SCs) in mouse and rat models of thoracic spinal cord injury (SCI), supporting a thoracic SCI clinical trial. Experts recommend in vivo preclinical testing of the intended clinical cell lot/line (CCL) in models with validity for the planned clinical target. We therefore tested the efficacy of two HuCNS-SC lines in cervical SCI: one RCL, and one CCL intended for use in the Pathway Study of cervical SCI in man. We assessed locomotor recovery and sensory function, as well as engraftment, migration, and fate. No evidence of efficacy of the CCL was observed; some data suggested a negative impact of the CCL on outcomes. These data raise questions about the development and validation of potency/comparability assays for clinical testing of cell products, and lack of US Food and Drug Administration requirements for in vivo testing of intended clinical cell lines.

INTRODUCTION

The first prospective study of spinal cord injury (SCI) prevalence in the USA revised the estimated number of individuals living with SCI upward by 5-fold to 1.3 million (Christopher and Dana Reeve Foundation, 2008). The average age at time of SCI is 34 years, resulting in a lifetime of paralysis associated with a host of medical complications. The impact of SCI in economic terms is highly disproportionate to the incidence of injury, rising to an average lifetime cost of several million dollars for individuals sustaining high-level cervical injuries. Critically, the majority of clinical SCI cases are at the cervical level (52.4%) (Christopher and Dana Reeve Foundation, 2008), making potential therapeutic interventions in cervical SCI rodent models a high priority.

Selection of a target clinical population is one key to translation of cell therapeutics. Two critical variables are treatment timing and vertebral level (thoracic versus cervical), both of which affect the incidence of spontaneous recovery in man (Fawcett et al., 2007). Rodent contusion models reproduce the principal pathophysiological features of clinical SCI with sensitive and relevant outcome measures (Stokes and Jakeman, 2002; Nishi et al., 2007). With respect to timing, the timeline of pathophysiological events following SCI in animal models versus the human condition is debatable; although many suggest that transplantation 9 days post-injury (DPI) in the rodent corresponds to the sub-acute clinical setting, while transplantation 30–60 DPI corresponds to the early chronic clinical setting (Houle and Tessler, 2003; Fawcett et al., 2007). Focusing enrollment for an SCI trial on chronic (>3 months post-SCI) cervical SCI subjects, compared with acute thoracic subjects, could reduce the enrollment required to attain statistical power to discriminate an improvement of 10 AIS (American Spinal Injury Association Impairment Scale) motor points dramatically from 250 to 25 AIS A subjects or from 1,100 to 50 AIS B subjects (Fawcett et al., 2007). Further, the larger pool of chronic SCI individuals may facilitate subject accrual, while an increased delay between injury and enrollment may improve the informed consent process (Anderson and Cummings, 2016).

With respect to vertebral level, there are compelling reasons to drive toward clinical trials focused on cervical SCI in more chronic cases. Many, however, have cautioned against proceeding to clinical trial for cervical SCI based on preclinical data in thoracic SCI models (Kwon et al., 2013). One reason for hesitation is increased recognition that cervical and thoracic injuries have a number of profound differences. For example, functional motor impairment across levels changes due to the anatomical characteristics of the spinal cord. Disruption of spinal circuitry due to systemic autonomic and immune effects is also level specific. Autonomic dysreflexias, particularly abnormal cardiovascular control, affect 50%–70% of human SCI patients with injury above T6 (Krassioukov and Claydon, 2006) but are rare when the injury is below this level. Accordingly, the impact of modulation of sprouting or connectivity via cell transplantation therapies could exert unanticipated effects in the case of high thoracic and

cervical SCI, which would not be evident in low thoracic SCI models. In parallel, disruption of descending sympathetic outflow specifically associated with cervical and high-level SCI has been shown to exert clinically significant and chronic splenic atrophy and immune suppression (Lucin et al., 2007; Zhang et al., 2013). Of note, SCI subjects in most cell therapeutic clinical trials would receive at least transient pharmacological immunosuppressive agents. In summary, these data demonstrate that injury level is a key variable in establishing not only efficacy but also safety in preclinical testing of investigational agents.

One therapeutic approach for SCI is cell transplantation. For this approach, cell survival is also a critical variable for preclinical models. The advantages of constitutively immunodeficient mice versus immunosuppression in immunocompetent mice in achieving maximal donor cell engraftment for xenotransplantation studies are significant (Anderson et al., 2011). This issue is particularly critical for establishing safety of a stem cell therapy, as tumor formation is impaired by a host immunorejection response (Dressel et al., 2008; Anderson et al., 2011). Accordingly, we have reported transplantation of multiple research-grade cell lines (RCLs) into both mouse and rat models of thoracic contusion SCI (Cummings et al., 2005; Hooshmand et al., 2009; Salazar et al., 2010; Piltti et al., 2013a, 2013b; Sontag et al., 2013, Sontag et al., 2014; Piltti et al., 2015). However, guidance from both the neurotransplantation and SCI fields emphasizes in vivo preclinical testing of the intended clinical cell lot/line (CCL) prior to proceeding in man. In the present study, we therefore tested the efficacy of human CNS-derived neural stem cell lines (HuCNS-SC) in an immunodeficient mouse model of cervical SCI. Two lines were tested: one derived as an RCL (HuCNS-SC RCL), the other intended for use in human cervical SCI under a funded NIH U01 (in what became the Pathway Study trial, NCT02163876). We assessed locomotor recovery and sensory function, as well as cell engraftment, migration, and neural lineage fate.

RESULTS

Transplantation of HuCNS-SC RCLs into a Unilateral Cervical SCI Model 9 DPI Results in Engraftment of Donor Human Cells and Functional Locomotor Recovery 12 Weeks Post-transplant

Based on our previously results transplanting multiple RCLs into both mouse and rat models of thoracic contusion SCI (Cummings et al., 2005; Hooshmand et al., 2009; Salazar et al., 2010; Piltti et al., 2013a, 2013b; Sontag et al., 2013, Sontag et al., 2014; Piltti et al., 2015), we investigated transplantation of the HuCNS-SC RCL at 9 DPI following unilateral cervical contusion SCI in Rag2γ mice. All mice received a unilateral right-sided contusion injury as described in the Experimental Procedures; accordingly, the right side of the cord is ipsilateral to the injury. A total of 75,000 HuCNS-SC RCLs were transplanted into four parenchymal sites (N = 12); mice in the control group received vehicle injections (N = 13). Bilateral transplantation was selected to parallel previous thoracic studies because of the potential for contralateral demyelination (Arvanian et al., 2009), and because HuCNS-SC transplanted only ipsilateral to the injury do not cross to the contralateral spinal cord. All of the Rag2γ mice receiving transplants exhibited engraftment, as identified by immunohistochemical staining for the human cytoplasm-specific antibody STEM121 (Figure 1A). Blinded, unbiased quantification of the number of STEM121$^+$/methyl green$^+$ cells using an optical fractionator probe revealed an average of 109,695 donor human cells/animal at 12 weeks post-transplant (WPT) (Figure 1B).

In locomotor assessment on a horizontal ladder beam task 12 WPT, mice receiving the HuCNS-SC RCL demonstrated a significant reduction in the number of ipsilateral forelimb errors compared with vehicle control (Figure 1C), as well as complete normalization of contralateral forelimb errors to pre-injury baseline (Figure 1D), demonstrating proof of concept for disease-modifying activity of HuCNS-SC RCL to improve locomotor function after cervical SCI.

Comparison of HuCNS-SC RCL and HuCNS-SC CCL Lines for Engraftment, Fate, Locomotor Recovery, and Sensory Parameters in a Unilateral Cervical SCI Model 60 DPI

We next sought to evaluate parameters important in establishing clinically relevant efficacy for cervical SCI: first, delayed transplantation, using a 60 DPI time point; second, in vivo comparability between the HuCNS-SC RCL employed above and the HuCNS-SC CCL intended for the human clinical trial in cervical SCI; third, a data profile that included assessment of allodynia and hyperalgesia and employed adult as opposed to aged mice. Groups were expanded to include injured only (no cell or vehicle injection) as an injection control and human fibroblast (hFB) transplant as a cellular control. Surgical and post-operative exclusions are described under Experimental Procedures and in Figures S1A–S1D.

RCL and CCL Engraftment in the 60 DPI Cohort

Engraftment data are shown for groups that received human cells (Figure 2A). All young adult Agouti Rag2γ(c) hybrid mice that received donor human RCL or CCL transplants exhibited engraftment. No sustained engraftment was observed in mice that received the hFB cellular control. Blinded, unbiased stereological quantification using the

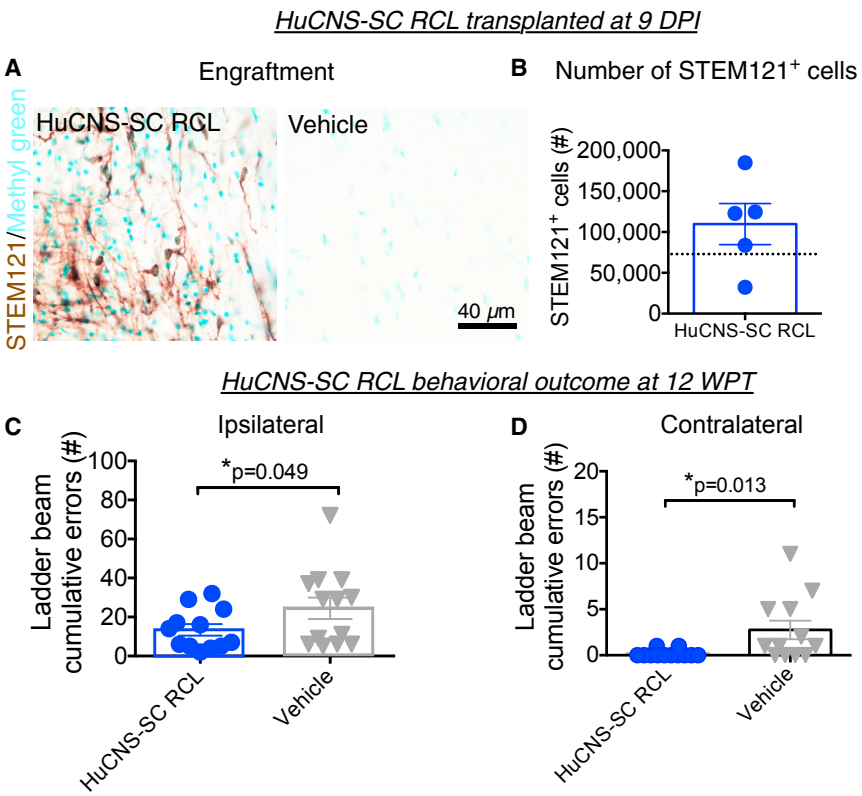


Figure 1. HuCNS-SC RCLs Exhibit Engraftment and Improve Locomotor Recovery in a 9 DPI Transplantation Paradigm

(A) Representative images from cervical SCI groups receiving HuCNS-SC RCL transplant (left) or vehicle control (right) 9 DPI. Animals were sacrificed 12 WPT. Sections were immunostained with a human-specific cytoplasmic marker STEM121 (brown) and counterstained with methyl green.

(B) Blinded, unbiased stereological quantification 12 WPT revealed an average of 109,695 ± 25,197 human cells in the RCL group (n = 9). Dashed line denotes transplant dose of 75,000 cells.

(C and D) Animals receiving RCL 9 DPI (n = 12) exhibited a significant decrease in horizontal ladder beam errors at 12 WPT for both ipsilateral (13.4 ± 3.0 versus 24.5 ± 5.5 for controls) and contralateral (0.2 ± 0.1 versus 3.8 ± 1.4 for controls) forepaws (Student's one-tailed t test, *p < 0.05) compared with controls (n = 13 ipsilateral, 12 contralateral). Data shown as means ± SEM.

optical fractionator probe revealed an average of 91,701 human cells/animal for the CCL and 200,754 human cells/animal for the RCL, which represented significantly greater engraftment of the RCL (Figure 2B; one-way ANOVA, $p < 0.0001$; post hoc Tukey's test of CCL versus RCL, $p < 0.0001$). However, the number of human cells was equivalent for the CCL 60 DPI cohort compared with the RCL 9 DPI proof-of-concept cohort data shown in Figure 1 (Figure 1B RCL 9 DPI 109,695 cells, Figure 2B CCL 60 DPI 91,701 cells; Student's two-tailed t test, $p > 0.4$), suggesting that there was adequate engraftment of the CCL to support recovery of function if the donor cells acted via the same or similar mechanisms. Stereology was used to analyze the migration of the CCL and RCL; no differences were observed in the rostral-caudal extent of migration between the two cell lines (Figure 2C).

RCL and CCL Fate in the 60 DPI Cohort

Cell fate data are shown only for groups that received HuCNS-SCs. We previously reported that multiple different HuCNS-SC RCLs exhibit robust differentiation along the oligodendroglial lineage after transplantation into rodent models of SCI 9, 30, or 60 DPI (Cummings et al., 2005; Hooshmand et al., 2009; Salazar et al., 2010; Piltti et al., 2013a, 2013b; Sontag et al., 2013). In contrast, these RCLs exhibit limited potential to generate oligodendroglial lineage cells in vitro (Sontag et al., 2013) and limited terminal differentiation into CC1-positive mature oligodendrocytes in the uninjured CNS in vivo (Sontag et al., 2014). Accordingly, we assessed the cell lineage fate of the RCL and CCL in this 60 DPI unilateral cervical model (Figure 3) in a random subset of animals from each group (N = 7/group) using double-labeling immunohistochemistry for $STEM121^+$ and doublecortin (DCX), nuclear Olig2, and APC/CC1, as described in the Experimental Procedures. Astroglial fate was determined by staining for the human-specific GFAP marker STEM123. Quantification was conducted via blinded, unbiased stereology using the optical fractionator probe in StereoInvestigator. All data are expressed as the proportion of $STEM121^+$/$marker^+$ donor human cells relative to total $STEM121^+$ cells in the same animal.

In contrast with previous studies in thoracic SCI at subacute and chronic time points, no evidence for neuronal lineage differentiation of donor cells, as evidenced by the lack of detection of $STEM121^+$/DCX^+ profiles, was observed (Figure 3A). These data were consistent with previous observations of the RCL transplanted at 9 DPI in this cervical SCI model (not shown), in which an extremely small percentage of DCX^+ neurons was observed and suggest that local cues in the cervical microenvironment are either

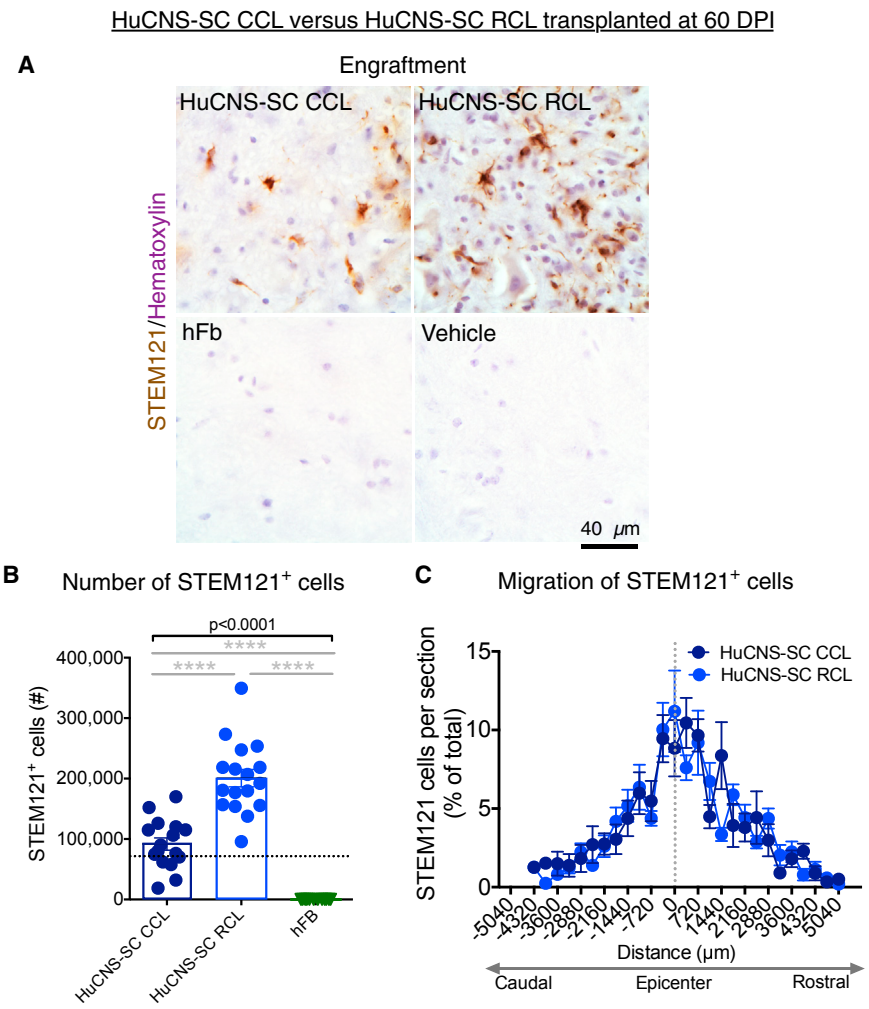


Figure 2. HuCNS-SC CCLs Exhibit Reduced Engraftment Compared with RCLs in a 60 DPI Paradigm

(A) Representative images from cervical SCI groups that received either a CCL, RCL, or hFb transplant, or vehicle control 60 DPI. Animals were sacrificed 12 WPT. Sections were immunostained with STEM121 (brown) and counterstained with hematoxylin (purple).

(B) Blinded quantification 12 WPT revealed an average of 91,701 human cells for the CCL group (n = 16), 200,754 for RCL group (n = 17), and no surviving human cells for the hFB group (n = 12) A one-way ANOVA revealed significant differences in survival (solid bar $p < 0.0001$; post hoc Tukey's test of CCL versus RCL and hFb, ****$p < 0.0001$). Dashed line denotes transplant dose of 75,000 cells.

(C) No difference was found in the rostral-caudal extent of human cell migration between the CCL (n = 16) and RCL (n = 17) at 12 WPT (unpaired t tests with Holm-Sidak multiple comparison correction, $p > 0.05$). Dashed vertical line indicates injury epicenter. Data shown as means ± SEM.

less permissive or fail to drive neuronal fate. Approximately 16% of CCL 60 DPI donor cells exhibited STEM123$^+$ immunolabeling, with no significant differences in proportional astroglial fate observed when RCL 60 DPI (23%) and CCL 60 DPI groups were compared (Figure 3B). This percentage was consistent with previous observations in thoracic SCI in which RCL were transplanted 30–60 DPI (Salazar et al., 2010; Piltti et al., 2013a, 2013b); in parallel, the majority of donor cells exhibited oligodendroglial lineage markers, regardless of group. However, there was a non-significant trend for a reduction in the STEM121$^+$/nuclear Olig2$^+$ proportion in CCL versus RCL cells (Figure 3C), and analysis of a mature oligodendroglial lineage marker (CC1) revealed approximately half the number of STEM121$^+$/CC1$^+$ donor human cells in CCL 60 DPI versus RCL 60 DPI transplanted animals (Figure 3D; 10% versus 18%, Student's two-tailed t test, $p < 0.007$). These data suggest that there were differences between the CCL and RCL in response to the injured microenvironment 60 DPI, which may have limited or delayed the generation of myelinating oligodendrocytes in CCL transplants.

Locomotor Recovery in the 60 DPI Cohort

No statistical differences were observed between injured only (no injection) and vehicle injection groups in any behavioral task; statistical comparisons shown were therefore based on comparison of groups receiving human cells versus vehicle injection. Recovery of function was assessed on four locomotor tasks: grip strength, cylinder reaching (percentage paw placement), horizontal ladder beam, and CatWalk step kinematic analysis. A summary of behavioral data is contained in Figure 4A. Although the RCL exhibited significant improvements in CatWalk Aa step pattern (coordination of alternating right-fore, right-hind, left-fore, left-hind) (Hamers et al., 2006) compared with the CCL at 12 WPT (Figure 4A; one-way ANOVA $p = 0.01$, post hoc Tukey's multiple comparison t test $p < 0.05$), there were no significant differences compared with vehicle controls.

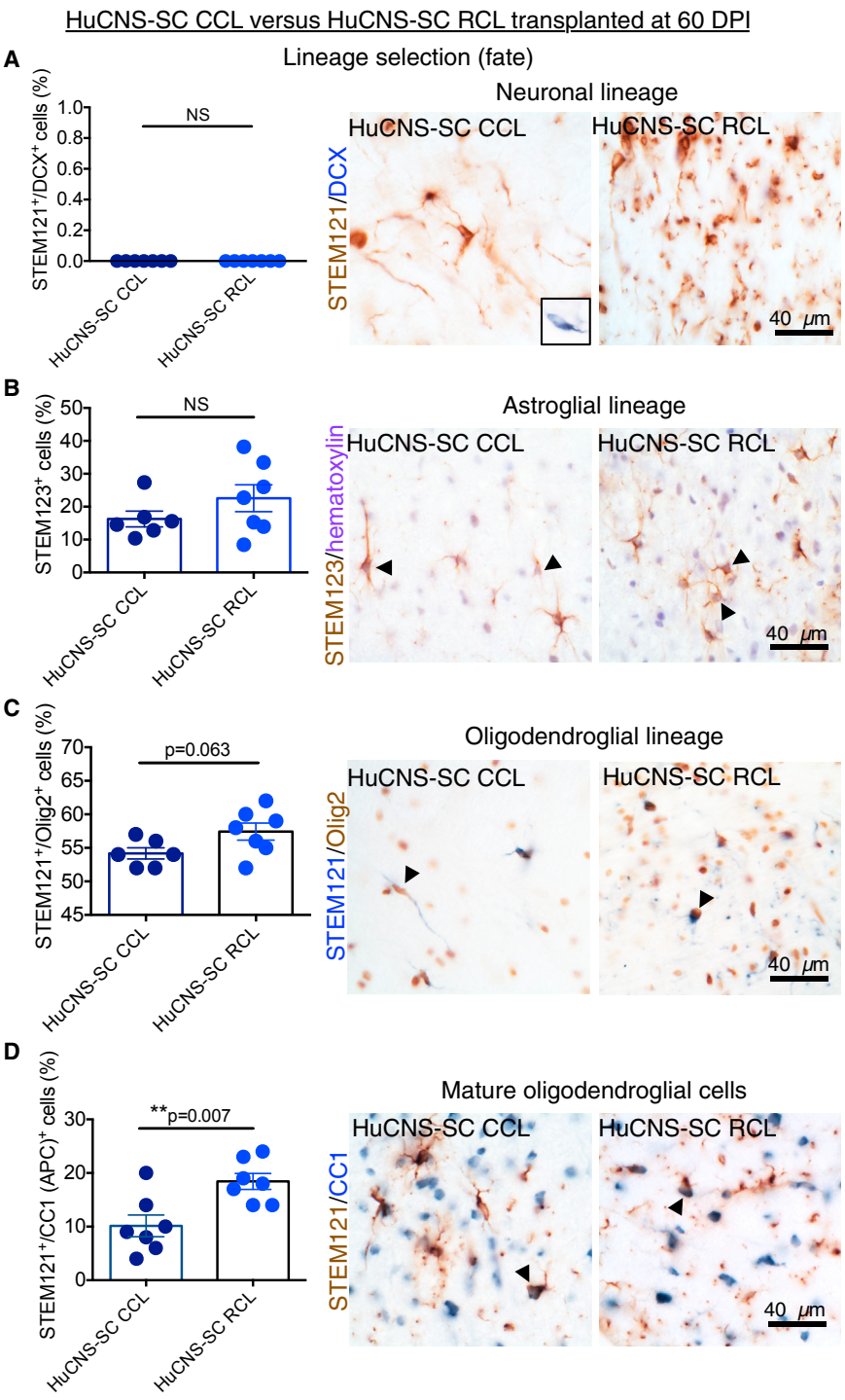


Figure 3. Lineage Analysis of HuCNS-SC CCL and RCL in the 60 DPI Paradigm

Data for proportional human cell fate were collected by blinded, unbiased stereology.

(A) Co-immunostaining for STEM121 (brown) and the early neuronal marker DCX (blue) revealed no STEM121$^+$/DCX$^+$ cells in either the CCL (n = 7) or RCL (n = 7) groups (Student's two-tailed t test, p > 0.05, n.s.). Inset shows positive control for DCX in the hippocampus.

(B) Immunostaining for human-specific GFAP (STEM123, brown) with hematoxylin counterstaining (purple) revealed no significant difference in STEM123 proportion between the CCL (n = 6) and RCL (n = 7) groups (Student's two-tailed t test, p > 0.2).

(C) Immunostaining for STEM121 (blue) and the oligodendroglial nuclear marker Olig2 (brown) revealed that the largest proportion of STEM121$^+$ cells were also nuclear Olig2$^+$ and there was a trend for a decrease in Olgi2$^+$ cells in CCL (n = 5) versus RCL (n = 7) transplants (Student's two-tailed t test, p = 0.06).

(D) Immunostaining for STEM121 (brown) and the mature oligodendroglial marker CC1 (blue) revealed a significant decrease in the proportion of STEM121$^+$/CC1$^+$ cells in CCL (n = 7) versus RCL (n = 7) transplants (Student's two-tailed t test, p < 0.007).

Arrowheads indicate double-positive cells. Data shown as means ± SEM.

No evidence for recovery of function was observed for the CCL in any of these tasks at 12 WPT or indeed at any time point (Figures 4 and S3). To test whether there was a relationship between the lack of observed locomotor recovery and 12 WPT engraftment, correlation analyses were performed for each of the principal locomotor measures collected for both the RCL and CCL. Although we have previously reported a reduction in hindlimb errors on the horizontal ladder beam that was correlated with increasing total STEM121$^+$ RCL engraftment after thoracic SCI (Hooshmand et al., 2009), we observed the opposite here; that is, an increase in right forelimb (ipsilateral) ladder

A

HuCNS-SC CCL versus HuCNS-SC RCL transplanted at 60 DPI

Behavioral outcome at 12 WPT

Behavioral measure	One-way ANOVA p-value	Transplantation groups				
		HuCNS-SC CCL	HuCNS-SC RCL	hFb	Vehicle	Injured
Cylinder RF	0.03 *	2.2±0.8	3.3±1.3	0.5±0.5 *	1.9±0.8	5.8±1.7
Cylinder LF	0.06	89.7±3.2	83.5±3.4	96.1±1.9	89.2±3.6	83.0±1.7
Grip strength RF	0.50	16.1±0.9	17.7±1.1	16.7±1.0	18.3±1.0	18.2±1.4
Grip strength LF	0.40	30.9±1.2	31.7±1.5	31.6±1.2	28.5±1.3	29.6±2.0
Ladder beam errors RF	1.00	7.1±0.9	7.3±1.0	7.3±0.8	6.8±0.6	6.4±1.2
Ladder beam errors LF	0.80	1.6±0.2	1.5±0.3	1.4±0.2	1.7±0.2	1.2±0.3
Catwalk swing speed	0.20	40.2±1.8	39.6±2.2	36.7±2.6	42.7±2.7	46.3±3.3
Catwalk swing speed LF	0.40	61.5±2.6	57.7±3.1	62.5±3.8	58.6±2.6	65.8±3.3
Catwalk duty cycle RF	0.40	44.4±2.0	43.4±1.6	40.4±4.4	46.8±1.7	45.1±1.8
Catwalk duty cycle LF	0.10	62.4±0.9	61.1±0.9	65.3±0.9	61.9±1.4	61.4±1.3
Catwalk BOS F	0.30	1.4±0.05	1.4±0.05	1.3±0.06	1.3±0.05	1.3±0.04
Catwalk regularity index	0.20	97.9±0.5	97.8±0.5	95.9±1.9	98.4±0.4	98.4±0.4
Catwalk run duration	0.40	2.8±0.1	3.0±0.1	3.2±0.1	2.6±0.2	2.6±0.2
Aa step pattern	0.01 **	0.2± 0.2	3.8±1.4 *	0±0.0	1.5±0.7	0.7±0.7
Ab step pattern	0.40	36.3± 7.8	30.5 ±4.1	32.1±	44.6± 7.9	28.8± 4.2
Ca step pattern	0.50	63.3± 7.8	59.0± 6.5	66.2±	50.6± 7.8	68.8± 4.5
Cb step pattern	0.30	0.0± 0.0	1.4± 1.0	0.3±0.3	0.9± 0.5	0± 0.0
Ra step pattern	1.00	0.0± 0.0	0.0± 0.0	0.0± 0.0	0.0± 0.0	0.0±0.0
Rb step pattern	0.10	0.0± 0.0	1.0±0.5	0.0± 0.0	0.7± 0.4	0.0±0.0

Relationship of human cell number and oligodendroglial fate to locomotor recovery in HuCNS-SC 60 DPI transplant groups

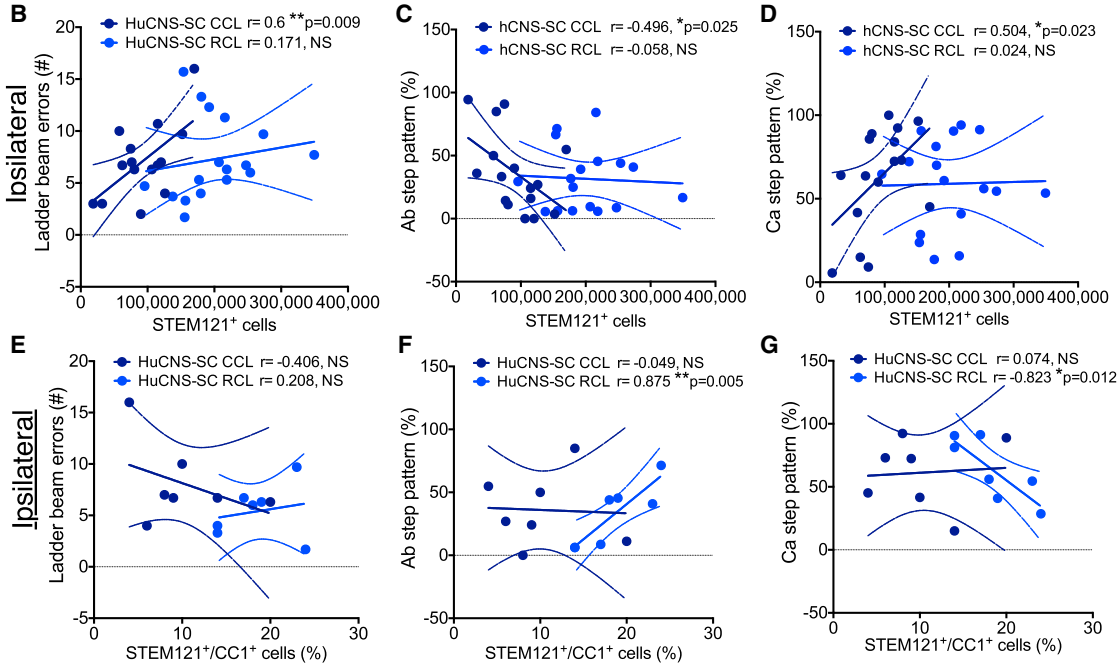


Figure 4. Analysis of Locomotor Recovery in HuCNS-SC 60 DPI Cohort Transplantation Groups versus Vehicle Controls 12 WPT

(A) Table showing average ± SEM values for cylinder reaching (percentage paw placement), grip strength, horizontal ladder beam errors, and CatWalk step kinematic analysis for each experimental group. RF, ipsilateral forelimb; LF, contralateral forelimb. Where one-way

(legend continued on next page)

errors was associated with an increase in the number of engrafted CCL cells (Figure 4B, dark blue circles; Pearson r = 0.6, p = 0.009). No correlation was observed between ipsilateral ladder errors and RCL engraftment (Figure 4B, light blue circles; Pearson r = 0.171, not significant [n.s.]). These data suggest a negative effect on locomotor recovery specific for the CCL, and failure of the RCL to produce robust locomotor recovery at this dose in young animals when transplantation was delayed to 60 DPI.

The finding of impaired function on the ladder beam was paralleled in kinematics analysis of gait on the CatWalk (Figures 4C and 4D). In young adult mice, the predominant pre-SCI step pattern was Ab (80%), while the predominant post-SCI step pattern shifted to Ca (60%). We focused on correlations of proportions of Ab and Ca patterns, anticipating that an improvement in locomotor recovery would be associated with an increase in Ab step pattern and a decrease in Ca step patterns. In contrast, CCL-treated animals exhibited a decrease in Ab step pattern and an increase in Ca step pattern in association with an increase in the number of engrafted CCL cells (Figure 4C, %Ab Pearson r = −0.496, p = 0.025; Figure 4D, %Ca Pearson r = 0.504, p = 0.023). No correlation was observed between Ab or Ca step pattern and RCL engraftment (Figure 4C, %Ab Pearson r = −0.0580, n.s.; Figure 4D, %Ca Pearson r = 0.0239, n.s.). Accordingly, these data suggest a negative effect on locomotor recovery specific for the CCL, and failure of the RCL to produce robust locomotor recovery in young animals when transplantation was delayed to 60 DPI.

Based on observed differences in $CC1^+$ oligodendroglial differentiation between the RCL and CCL, we also conducted correlation analyses for ladder beam errors, Ab step pattern, and Ca step pattern for these cell lines. No significant relationships were found between ladder beam errors and proportion of $STEM121^+/CC1^+$ cells for either the RCL or CCL (Figure 4E, RCL Pearson r = 0.208, n.s.; CCL Pearson r = −0.406, n.s.). Similarly, the CCL exhibited no relationship between proportion of $STEM121^+/CC1^+$ cells for either Ab or Ca step patterns (Figure 4F, Ab Pearson r = −0.049, n.s.; Figure 4G Ca Pearson r = −0.074, n.s.). However, analysis of these relationships for the RCL was significant (Figure 4F, Ab Pearson r = −0.875, p = 0.005; Figure 4G, Ca Pearson r = −0.823, p = 0.012), suggesting that the potential to exert donor-cell-driven repair in chronic cervical SCI may be associated with in vivo capacity to generate terminally differentiated oligodendrocytes.

Sensory Assessment in the 60 DPI Cohort

A significant concern is the potential for cell engraftment, or specific lineage selection, to induce or exacerbate neuropathic pain syndromes in SCI (Hofstetter et al., 2005; Macias et al., 2006). Accordingly, animals were assessed for mechanical allodynia using Von Frey testing and hyperalgesia using Hargreaves testing. Critically, although reduced, allodynia/hyperalgesia can be detected in constitutively immunodeficient animals (Moalem et al., 2004; Kleinschnitz et al., 2006). No changes were observed between any groups in these measures either immediately prior to sacrifice at 12 WPT or in two-way ANOVA across time (Figure S4), suggesting neither a cell-based impairment of sensory function nor initiation of a neuropathic pain syndrome.

Evaluation of the CCL for Engraftment, Fate, Locomotor Recovery, and Sensory Parameters in a Unilateral Cervical SCI Model 9 DPI

While RCL transplantation into the cervical spinal cord 9 DPI resulted in recovery of locomotor function in aged Rag2γ mice, the effect of RCL transplantation into the cervical spinal cord 60 DPI in young Agouti Rag2γ(c) hybrid mice was attenuated, and correlational analyses suggest CCL transplantation into this model produced some decrements in function. Together, the data suggest that less benefit was achieved after transplantation into a delayed/chronic cervical SCI paradigm, and that the RCL and CCL exerted different effects after cervical SCI. To investigate this variation, we evaluated cell engraftment, fate, locomotor recovery, and sensory function following CCL transplantation into the cervical spinal cord 9 DPI in a separate cohort of young Agouti Rag2γ(c) hybrid mice.

CCL Engraftment in the 9 DPI Cohort

All Agouti Rag2γ(c) hybrid mice receiving CCL at 9 DPI exhibited engraftment 12 WPT, while no sustained engraftment was observed in mice receiving the hFB cellular control (Figure 5A). Blinded, unbiased stereological quantification revealed an average of 118,757 human cells/animal (Figure 5B), which was significantly greater compared

ANOVA reached significance (*$p \leq 0.05$, **$p \leq 0.01$), Tukey's multiple comparison t tests were conducted. Significant differences were found in ipsilateral cylinder reaching between the hFB and Injured only groups (*$p \leq 0.05$), and in Aa step pattern between the RCL, CCL, and hFB groups (*$p \leq 0.05$). n = 15–16 for CCL, n = 16–17 for RCL, n = 11–12 for hFB, n = 15–17 for Vehicle, and n = 9–10 for Injured only (see Figure S1D for exact numbers). BOS, base of support.

(B–G) Pearson correlations were conducted between the number of human $STEM121^+$ cells and (B) ipsilateral forelimb ladder beam errors, (C) Ab step pattern, and (D) Ca step pattern. Pearson correlation were also conducted between $STEM121^+/CC1^+$ human oligodendroglial cell proportion and (E) ipsilateral forelimb ladder beam errors, (F) Ab step pattern, and (G) Ca step pattern. Dashed lines indicate confidence intervals of 95%. n = 7 for CCL and n = 7 for RCL in CC1+ correlational analyses.

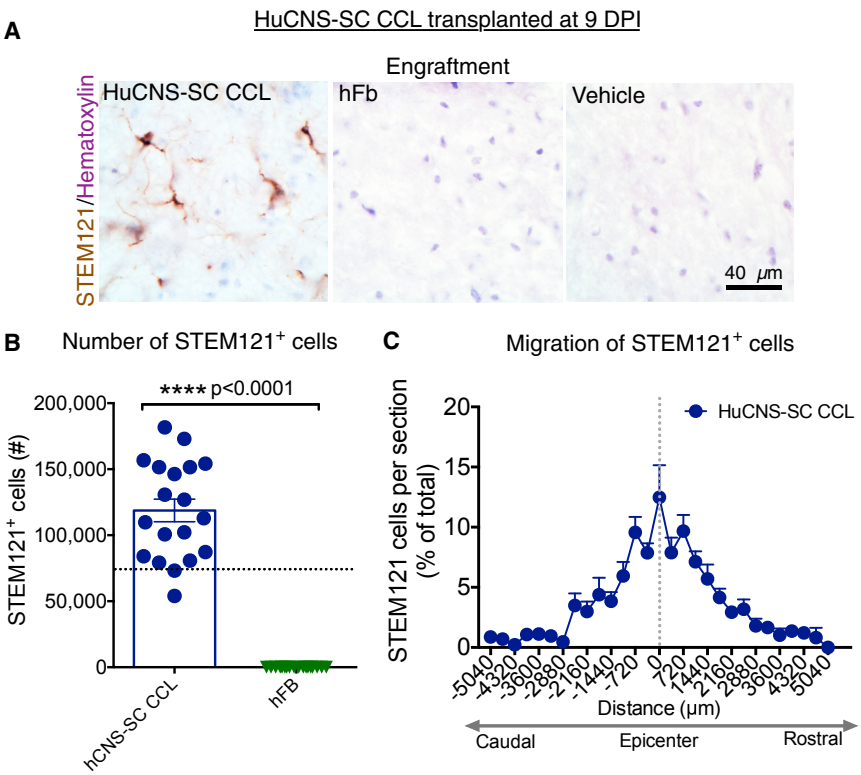


Figure 5. HuCNS-SC CCL Exhibit Engraftment and Rostral-Caudal Migration in a 9 DPI Paradigm

(A) Representative images from cervical SCI groups that received CCL or hFb transplant, or vehicle control 9 DPI. Animals were sacrificed 12 WPT. Sections were immunostained with STEM121 (brown) and counterstained with hematoxylin.

(B) Blinded, unbiased quantification 12 WPT revealed an average of 118,757 human cells for the CCL group (n = 19) and no human cells for the hFB group (n = 18) (Student's two-tailed t test, $p < 0.0001$). Dashed line indicates the transplant dose of 75,000 cells.

(C) Rostral-caudal extent of human cell migration for the CCL at 12 WPT. Dashed vertical line indicates injury epicenter (n = 19). Data shown as means ± SEM.

with the 60 DPI CCL cohort Figure 2B; a comparison is shown in Figure S2A (18,757 versus 91,701, Student's two-tailed t test, $p < 0.05$). This contrasts with previous observations in thoracic SCI, where we observe no significant differences in surviving cells quantified 12–16 WPT after human neural stem cells transplantation at 0, 9, or 30 DPI. However, comparison of migration between the 9 (Figure 5C) and 60 DPI (Figure 2C) CCL cohorts revealed no differences in the rostral-caudal extent of migration at any distance from the lesion epicenter (unpaired t test with Holm-Sidak multiple comparison correction, n.s.).

CCL Fate in the 9 DPI Cohort

Analysis of human cell fate was conducted as described for 60 DPI cohorts. As for the 60 DPI CCL cohort, no evidence of neuronal lineage differentiation was observed (Figure 6A). Approximately 33% of human CCL 9 DPI were positive for the human astroglial marker STEM123 (Figure 6B). In parallel with the 60 DPI cohort, the largest proportion of human cells exhibited oligodendroglial lineage markers and were positive for either STEM121$^+$/Olig2$^+$ (Figure 6C; 55%) or STEM121$^+$/CC1$^+$ (Figure 6D; 10%). No differences were observed between the 9 DPI and 60 DPI CCL cohorts in neuronal or oligodendroglial lineage fate proportions (Figure S2B, Student's two-tailed t test, $p > 0.05$). However, comparison of the CCL 9 DPI and 60 DPI cohorts revealed a doubling of STEM123$^+$ astroglial cells in the 9 DPI cohort (Figure S2B, Student's two-tailed t test, $p < 0.009$). Differences in the post-injury cervical SCI microenvironment associated with transplantation timing and/or variation between CCL shipments, corresponding to the 9 DPI and 60 DPI cohort transplantation dates, may have influenced CCL fate. In comparison, previous experience with RCL has not demonstrated an increase in astroglial fate in mice receiving human cells at 30 versus 9 DPI (Cummings et al., 2005; Salazar et al., 2010).

Locomotor Recovery in the 9 DPI Cohort

In contrast to the 9 DPI RCL cohort, no evidence for recovery of function was observed for the 9 DPI CCL cohort in any locomotor assessment at 12 WPT or any time point (Figures 7 and S5). A summary of behavioral data is shown in Figure 7A. Analysis of locomotor recovery versus engraftment at 12 WPT did not reveal significant correlations between the number of engrafted STEM121$^+$ CCL cells and right forelimb (ipsilateral) errors on the horizontal ladder beam (Figure 7A; Pearson $r = -0.218$, n.s.) or in any CatWalk step patterns, including Ab (Figure 7B; Pearson $r = 0.147$, n.s.) or Ca (Figure 7C; Pearson $r = -0.068$, n.s.). Similarly, analysis of locomotor recovery versus fate did not reveal significant correlations between the number of engrafted STEM121$^+$/CC1$^+$ CCL cells and right forelimb (ipsilateral) errors on the horizontal ladder beam (Figure 7D; Pearson $r = 0.278$, n.s.) or in any CatWalk step

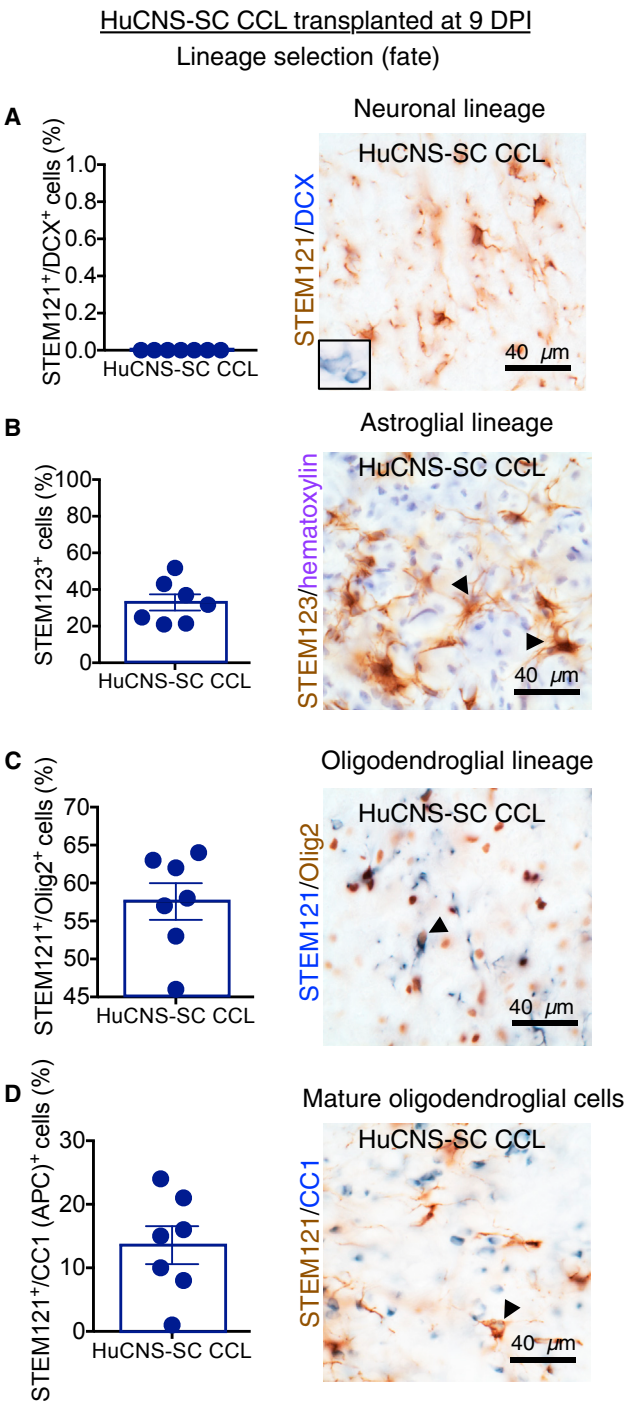


Figure 6. Lineage Analysis of HuCNS-SC CCL in the 9 DPI Paradigm

Data for proportional human cell fate were collected by blinded, unbiased stereology

(A) Immunostaining for STEM121 (brown) and DCX (blue) revealed no STEM121$^+$/DCX$^+$ cells in CCL 9 DPI animals. Inset shows positive control for DCX in cortex.

(B) Immunostaining for human-specific GFAP (STEM123, brown) with hematoxylin counterstaining (purple) revealed 33% ± 4.4% of CCL 9 DPI were positive for this astroglial marker.

(C) Immunostaining for STEM121 (blue) and Olig2 (brown) revealed that the largest proportion of STEM121$^+$ cells in the 9 DPI transplant group were positive for this oligodendroglial marker (57.6% ± 2.4%).

(D) Immunostaining for STEM121 (brown) and CC1 (blue) revealed 13.6% ± 3% of CCL 9 DPI were positive for this mature oligodendroglial marker.

Arrowheads indicate double-positive cells. n = 7 for all lineage tests. Data shown as means ± SEM.

patterns, including Ab (Figure 7E; Pearson r = 0.143, n.s.) or Ca (Figure 7F; Pearson r = −0.366, n.s.). Together, these data suggest that the CCL failed to promote functional locomotor recovery in this animal model of cervical SCI; however, no significant decrements in function were observed on correlational analyses.

Sensory Parameter Assessment in the 9 DPI Cohort

All animals were assessed for mechanical allodynia using Von Frey testing and hyperalgesia via Hargreaves testing. No changes were observed between any groups in either two-way ANOVA across time or immediately prior to sacrifice 12 WPT (Figure S6). Accordingly, these data suggest that CCL transplantation 9 DPI neither altered sensory function nor initiated a neuropathic pain syndrome.

DISCUSSION

HuCNS-SC CCLs have been tested in human clinical trials in several paradigms. An HuCNS-SC line was authorized by the US Food and Drug Administration (FDA) for testing in the lysosomal storage disorder neuronal ceroid lipofuscinosis (NCL) (NCT00337636). Six patients were transplanted and the study was completed in 2009. The results suggested that the direct transplantation of HuCNS-SCs into the CNS was safe (Selden et al., 2013). An HuCNS-SC line was also authorized for testing in the lethal disorder Pelizaeus-Merzbacher disease (NCT01005004). Four patients received this HuCNS-SC line and the study was completed in 2012 (Gupta et al., 2012), again supporting a favorable safety profile. In 2010, an HuCNS-SC line was authorized by Swissmedic for a phase I/II trial in thoracic SCI (NCT01321333) in Zurich (Health Canada and the FDA later adding two North American sites). An HuCNS-SC line was transplanted into 12 SCI patients; the trial was completed in April 2015. Interim data were presented at the fourth Joint International Spinal Cord Society and American Spinal Injury Association meeting in Montreal (May 14, 2015). However, as of September 30, 2016, final results and/or a peer-reviewed publication of this study has yet to be published. The principal objective of the current study was to test the efficacy of the HuCNS-SC CCL intended for use in the human cervical SCI Pathway Study

A

HuCNS-SC CCL transplanted at 9 DPI

Behavioral outcome at 12 WPT

Behavioral measure	One-way ANOVA p-value	Transplantation groups			
		HuCNS-SC CCL	hFb	Vehicle	Injured
Cylinder RF	0.09	0.6± 0.6	1.5± 1.1	4.4± 1.4	2.9±1.3
Cylinder LF	1.0	76.2± 6.1	80.0± 7.0	79.1± 3.8	77.3±9.
Grip strength RF	0.20	19.1± 0.9	16.2± 1.1	18.1± 1.0	19.0±1.
Grip strength LF	0.90	39.0± 2.1	41.1± 2.3	39.3± 2.1	38.2±2.
Ladder beam errors RF	0.06	6.9± 0.9	8.3± 0.6*	5.4± 0.7	6.1±0.7
Ladder beam errors LF	0.06	1.9± 0.3*	1.5± 0.2	1.3± 0.2	0.8± 0.2
Catwalk swing speed RF	0.30	46± 2.7	42.2± 2.2	49.0± 3.0	46.8±2.
Catwalk swing speed LF	0.30	74.8± 4.0	67.4± 2.1	67.0± 4.3	67.0±4.
Catwalk duty cycle RF	0.80	38.4± 2.1	38.0± 1.4	39.9± 2.3	40.5±2.
Catwalk duty cycle LF	0.09	62.3± 1.1	61.3± 1.1	59.6± 1.3	58.0±0.
Catwalk BOS F	0.30	1.4± 0.04	1.4± 0.05	1.3± 0.05	1.3±0.0
Catwalk regularity index	0.20	94.1± 1.5	96.0± 1.3	94.2± 1.8	99.2±0.
Catwalk run duration	0.90	2.4± 0.1	2.3± 0.1	2.2± 0.1	2.4±0.2
Aa step pattern	0.4	1.2± 0.6	1.6± 0.9	2.3± 1.0	0.0±0.0
Ab step pattern	0.80	30.7± 5.6	29.8± 5.8	34.5± 6.5	38.0±7.
Ca step pattern	1.0	61.5± 5.8	63.2± 6.0	58.9± 5.9	61.5±7.
Cb step pattern	0.30	0.6± 0.4	0.6± 0.4	0.0± 0.0	0.0± 0.0
Ra step pattern	0.30	1.5± 0.8	0.4± 0.3	0.5± 0.3	0.0± 0.0
Rb step pattern	0.40	1.1± 0.5	0.7± 0.4	1.3± 0.6	0.0± 0.0

Relationship of human cell number and oligodendroglial fate to locomotor recovery in HuCNS-SC CCL transplanted at 9DPI

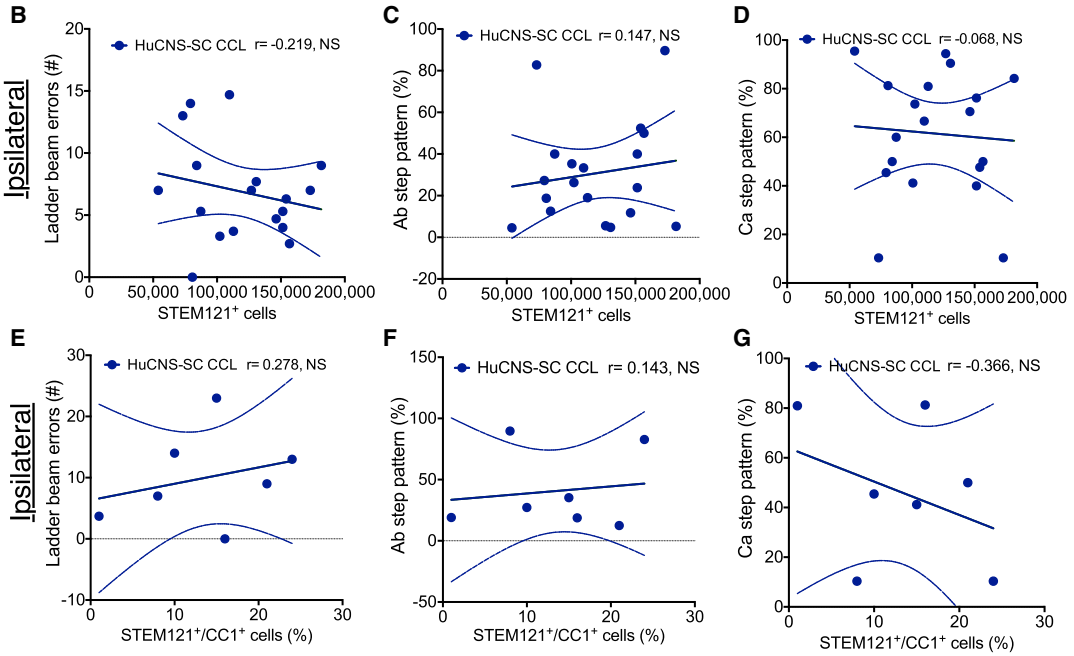


Figure 7. Locomotor Recovery in HuCNS-SC 9 DPI Cohort Transplantation Groups versus Vehicle Controls 12 WPT

Data for proportional human cell fate were collected by blinded, unbiased stereology.

(A) Average ± SEM values for cylinder reaching (percentage paw placement), grip strength, horizontal ladder beam errors, and CatWalk kinematic analysis for each group. RF, ipsilateral forelimb; LF, contralateral forelimb. No significant differences were found between groups

(legend continued on next page)

under an NIH-funded U01 in a unilateral cervical contusion injury model, assessing locomotor recovery and sensory function, as well as cell engraftment, migration, and neural lineage fate.

Although we have previously demonstrated evidence for recovery of motor function in both thoracic and cervical models using several different research cell lines (RCL), no evidence in support of the efficacy of the tested intended HuCNS-SC CCL was detected in the current cervical SCI study. In contrast, correlative measures of total donor cell number and locomotor function suggested, in some cases, a negative impact of engraftment of the CCL on functional outcome in this animal model. A clinical trial testing this HuCNS-SC CCL was initiated in December 2014 for cervical SCI (NCT02163876) after preliminary analysis of the dataset in the present report was reported to StemCells Inc. A press release of interim 6-month data for the trial (November 18, 2015) reported improvements in motor strength in 4/5 subjects, in contrast to the animal data presented here. Subsequently, citing a lack of significant improvements and the lack of a trend for improvements over time, StemCells Inc. terminated the Pathway Study on May 31, 2016.

These facts raise several issues for stem cell transplantation. There has been extensive discussion regarding the validity of animal models in clinical translation (van der Worp et al., 2010), based on the failure of animal models of disease to predict efficacy in the clinical trial setting. Publication bias and overprediction of efficacy is suggested to account for as much as a third of this discrepancy (Sena et al., 2010), and methodological/RIGOR flaws in animal studies have also been raised as significant concerns (Lapchak, 2012; Chang et al., 2015). It can be difficult to determine the applicability of endpoints in an animal model to the human condition, or the timing of a therapeutic in rodents versus timing in man (Henderson et al., 2013). Conversely, poorly designed clinical studies may play a role in overestimating or underestimating clinical impact (Moller, 2014). Accordingly, publication of preclinical data from carefully controlled and properly blinded animal studies is key for advancing the field. Multiple RCLs have previously been shown to improve recovery of function in multiple thoracic SCI models. In addition, the RCL employed in the proof-of-concept cohort here did demonstrate efficacy in this cervical SCI model. Moreover, in our preclinical animal studies, all RIGOR recommendations and other standards were met. One interpretation of these data is that these observations derive from variation between cell lines and/or in cell manufacture/processing and not in the model itself or in the experimental execution of the model.

For stem cell therapies, generation of cell lots and lines with consistent potency and comparability is recognized by the field and the FDA as a significant issue for clinical translation (Hyun et al., 2008; Lo et al., 2008). Inconsistencies in scale-up and production of good manufacturing practice (GMP) CCLs, failure to develop potency/comparability assays based on demonstrable clinically significant activity rather than simplicity, and failure of potency/comparability assays to offer adequate analysis of clinically significant endpoints are issues that have been addressed extensively in the context of mesenchymal stem cell products (MSCs). MSCs represent an exemplar in which the failure of pivotal clinical trials has been linked to these factors (Galipeau, 2013; Chinnadurai et al., 2015). Accordingly, in addition to controversies over the validity of animal models and quality of preclinical design, an additional variable may be the development and validation of potency/comparability assays for clinical testing of cell products, including the lack of FDA requirements for in vivo testing. Critically, failure to address and define the source of variation between research, process development, and clinical cell lots and lines, or to conduct in vivo testing of all cell product lots and lines used for clinical transplantation, is an issue for not only efficacy, but safety, as the in vivo factors controlling donor differentiation, cell division, and tumorigenesis remain poorly defined, and a focus on in vitro assays that are poorly linked to efficacy may fail to detect critical variations that result in an altered risk profile. We suggest that this raises concerns regarding both the adequacy of current standards for demonstration of potency and comparability between therapeutic cell lots and lines, as well as a potential issue for informed consent during patient enrollment (Anderson and Cummings, 2016).

using one-way ANOVA $p = 0.06$; however Tukey's multiple comparison revealed significant an increase in ipsilateral forelimb ladder beam errors in the hFB group compared with vehicle group ($p < 0.05$) and in contralateral forelimb ladder beam errors in CCL 9 DPI group compared with injured only group ($p < 0.05$), denoted by *. n = 18–19 for CCL, n = 17–18 for hFB, n = 18–20 for Vehicle, and n = 9–10 for Injured only (see Figure S1D for exact numbers).

(B–G) Pearson correlations were conducted between the number of human STEM121$^+$ cells and (B) ipsilateral forelimb ladder beam error number, (C) Ab step pattern, and (D) Ca step pattern (n = 18). Pearson correlations were also run between STEM121$^+$/CC1$^+$ human oligodendroglial cell proportion and (E) ipsilateral forelimb ladder beam errors, (F) Ab step pattern, and (G) Ca step pattern (n = 7). No significant differences were observed, suggesting neither improvements nor decrements in function at 9 DPI. Dashed lines indicate confidence intervals of 95%.

In response to these data, StemCells Inc. noted that the CCL cells tested herein are not the cells that were tested in the Pathway Study, because cells were sent to the University of California, Irvine (UCI) from the "process development laboratory" and not produced under current GMP conditions/manufacturing requirements of a clinical trial. StemCell Inc.'s response is contrary to the milestones of the NIH-funded U01 for testing of the "intended CCL". However, we do not have access to the safety/toxicology profile submitted to gain FDA authorization for the trial nor the final clinical product administered in the Pathway Study. If one accepts that the CCL cells used herein do not share sufficient comparability with the final clinical product to be considered as "representative", then we believe the Pathway Study went forward in the absence of in vivo efficacy testing. Conversely, if one views the "process development laboratory" and final clinical product as substantially similar, then we would argue that the Pathway Study went forward with cells that failed to yield preclinical efficacy. Many scientists are unaware that in vivo preclinical testing of the final clinical product is not required by the FDA. As we have noted (Anderson and Cummings, 2016), FDA guidance states that because "human-derived cellular therapy products intended for clinical administration in animals may not be informative" (due to the species specific nature of some paradigms or products), "testing of an analogous product may be a suitable alternative" (Center for Biologics Evaluation and Research, 2013).

Of course, when the final analysis and follow-up are completed for the Pathway Study, the key issue will be whether the final clinical product demonstrated a positive or negative safety and tolerability profile, as well as preliminary evidence related to efficacy outcome measures in humans. If negative safety data or a failure in efficacy are ultimately observed in the clinical trial, it is likely that hindsight will call out the failure to match preclinical and clinical studies in terms of the tested product. Conversely, if positive safety and/or efficacy profiles were to have been observed in the clinical trial, one might have asserted that preclinical models of cervical SCI in rodents are not predictive, and/or that the details of cell manufacturing significantly affect outcome in both animals and humans. In the end, we have no way to conduct an objective assessment of preclinical versus clinical discrepancies, which raises a more general problem for translational research supported by academic-industry partnerships.

Are the cells used in the NCL trial, the Pelizaeus-Merzbacher Disease trial, or the Zurich thoracic SCI trial derived from the same donor or by the same manufacturing process as those used in the Pathway Study? Because all cell lines, whether from different donors or different manufacturing preparations, have been designated as "HuCNS-SC" by StemCells Inc., it is impossible for a subject enrolling in a trial using this product to fully understand the basis on which the trial was founded, and come to an informed individual decision on participation. We suggest that open disclosure of specific stem cell product designations (e.g., standardized clinical cell line reference numbers), should be required, negotiated as a part of academic-industry collaborations, as well as for publication of preclinical and clinical data, and institutional review board approval of consent documents. This is a standard that we are unable to meet due to nondisclosure restrictions; accordingly, multiple different cell lines are referred to herein by the common appellation HuCNS-SC. For clarity in this regard, it should be stated that the CCL referenced in this article is not the same line as the CCL reported on by Marsh et al. (2017), or any cell line previously reported on by our laboratory.

Finally, it is important to address the relationship between translational/preclinical research and clinical trial success rate (Roberts et al., 2012; Perrin, 2014). Although many factors may contribute, including lack of alignment between animal models and human disease, the lack of correspondence between preclinical research and clinical trial success may also suggest a need for conducting sufficient basic research to understand the mechanism(s) of action (MOA) of a particular cell therapy and enable potency/comparability assay design that can robustly detect variation between cell lots and lines. Until optimized in vitro assays are available, this may require in vivo testing of final clinical products that have completed full-release testing. It has been argued that the MOA should not be required in order to proceed with testing a drug or cell therapy in man. This position posits that were we to wait until every aspect of a particular cell line were understood, we would never proceed to clinical testing. In fact, with respect to evidentiary standards for drug testing and approval, the prevailing opinion is that "theories about MOA of a drug or disease mechanisms play important parts in drug development and approval, but they are entirely subsidiary to the fundamental questions that must be answered in the course of drug approval; namely, is a drug effective, and is it safe in use." (Katz, 2004). While this perspective has a logical degree of practicality, we suggest that the failure (or disincentive) to understand MOA may be an alternative reason for failure in translational medicine and clinical trials. Finally, since the Pathway Study failed to show efficacy in humans, we will never know if this failure was because neural stem cells, in general, are not effective for human SCI or, rather, that the wrong cell line was tested prematurely in humans.

EXPERIMENTAL PROCEDURES

All studies were in accordance with the Institutional Animal Care and Use Committee and Human Stem Cell Research Oversight at UCI. All data were maintained under good laboratory

practice-like protocols, with an assigned data monitor. Animal care, behavior, and analysis were performed by investigators blinded to group, and random group allotment was used. Inclusion/exclusion criteria were established in advance (see Supplemental Experimental Procedures). Additional detail on experimental procedures, randomization, blinding, exclusions, cells, injuries and surgeries, transplantation, and analyses is contained in Supplemental Experimental Procedures.

HuCNS-SC and hFb

Sorted HuCNS-SC lines were provided by StemCells Inc.; details about cells are described in the Supplemental Experimental Procedures. RCL and CCL cells were shipped overnight to UCI. Fresh RCL or CCL shipments were received each surgery day. Cell yield, viability, and preparation data were recorded for each vial received (Table S1). Human mesenchymal stromal cell hFbs (Cell Applications) were thawed and cultured at UCI in DMEM with 10% fetal bovine serum and glutamine for 7 days prior to transplantation.

Contusion Injuries

For the proof-of-concept cohort, we used 18-month-old, commercially available, Rag2γ(c) female mice (Taconic). For the main study, 10 to 12 week old female Agouti Rag2γ(c) hybrid mice were used (StemCells Inc.). Animals were anesthetized, spinal cords at C5 vertebral level were exposed by laminectomy, stabilized, and unilateral 30-kDa contusion injuries with 5 s dwell time were administrated with a 1 mm diameter tip using an IH Impactor (Precision Systems & Instrumentation). Animals received standard post-operative care, including buprenorphine, lactated Ringer's, bladder care, and antibiotics.

Transplantation

Mice were anesthetized 9 or 60 DPI and a total volume of 1 μL (250 nL per injection site) of cell suspension (75,000 cells) or vehicle was injected via two rostral injections and two caudal injections 0.75 mm from midline.

Exclusions

Thirty-three animals entered the proof-of-concept study cohort; 25 completed the study. A total of 147 animals entered the main study cohorts; 139 completed the study. There were no animal exclusions due to engraftment failure or histological issues. Pre- and post-injury animal exclusions, Grubbs exclusions, and final numbers for statistical analysis are detailed in Figures S1A–S1D.

Assessment of Locomotor and Sensory Function

All tasks were assessed prior to injury (baseline) and at 3, 8, and 12 WPT. The horizontal ladder beam task was performed as described (Salazar et al., 2010) using CatWalk XT (Noldus v9.0). Forelimb-use asymmetry was assessed using a cylinder task. Forepaw grip strength was measured for each paw alone and together in five trials per mouse using a Dunnett-style grip strength meter (Pawar et al., 2015). Mechanical allodynia was assessed using a Von Frey test (Salazar et al., 2010). Thermal hyperalgesia was assessed using a Hargreaves test (Piltti et al., 2013a, 2013b).

Histology and Stereological Quantification

At 12 WPT, mice were terminally anesthetized, and perfused, and cord segments were dissected (C1–T2 roots), post-fixed, cryoprotected with 20% sucrose, and flash frozen (Hooshmand et al., 2009). Coronal 30 μm sections were cryostated and immunostained. The primary antibodies and secondary antibodies used are listed in Table S2. Final animal numbers for analysis are listed in Figure S1. Quantification was via unbiased stereology. Migration of human cells was analyzed as percentage of the cells per section relative to total number of human cells. A random set of animals were selected for proportional counting of cell fate (N = 7/group).

Statistics

All data are means ± SEM; statistics were performed using Prism v6 (GraphPad Software). Comparisons between groups were analyzed using one-way ANOVA combined with Tukey's post hoc t tests or Student's one or two-tailed t tests. Locomotor and sensory function were compared using two-way repeated measures ANOVA combined with multiple comparison corrected/Bonferroni post hoc t tests. Migration was analyzed using unpaired two-tailed t tests with Holm-Sidak multiple comparison correction. Correlation between human cells and ladder beam errors or percentage of step patterns were assessed using a Pearson correlation coefficient. A p value of ≤ 0.05 was considered to be significant. In the figure legends, n denotes the number of individual mice.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2016.12.018.

AUTHOR CONTRIBUTIONS

A.J.A. designed experiments, wrote the manuscript, trained and supervised staff, and contributed to data analysis and finalization of figures for publication. K.M.P. supervised and conducted cell preparation for transplantation, prepared figures for publication, and contributed to data analysis and to manuscript writing. M.J.H. supervised cell transplantation, maintained experimental blinding for surgeries and behavioral analyses, and contributed to data analysis and manuscript writing. R.A.N. supervised all surgeries and behavioral testing, and contributed to experimental planning and data analysis. B.J.C. designed experiments, and contributed to final data analysis, preparation of figures for publication, and manuscript writing. K.M.P., M.J.H., and R.A.N. contributed equally.

ACKNOWLEDGMENTS

We thank staff of the Christopher and Dana Reeve Foundation Core (CDRF core), especially B. Ott, J. Lucero, O. Mendez, H. Liu, and C. Worne. This study was supported by an NIH grant (U01-NS079420), the CDRF (AAC-2005), and a canceled Sponsored Research Agreement (STEM-1000643) from StemCells Inc. This work was supported, in part, by a Sponsored Research Agreement (STEM-1000634) from StemCells Inc. to A.J.A. and B.J.C. However,

this Sponsored Research Agreement was canceled during the collection/analysis of the data.

Received: June 22, 2016
Revised: December 14, 2016
Accepted: December 19, 2016
Published: February 14, 2017

REFERENCES

Anderson, A.J., and Cummings, B.J. (2016). Achieving informed consent for cellular therapies: a preclinical translational research perspective on regulations versus a dose of reality. J. Law Med. Ethics *44*, 394–401.

Anderson, A.J., Haus, D.L., Hooshmand, M.J., Perez, H., Sontag, C.J., and Cummings, B.J. (2011). Achieving stable human stem cell engraftment and survival in the CNS: is the future of regenerative medicine immunodeficient? Regen. Med. *6*, 367–406.

Arvanian, V.L., Schnell, L., Lou, L., Golshani, R., Hunanyan, A., Ghosh, A., Pearse, D.D., Robinson, J.K., Schwab, M.E., Fawcett, J.W., and Mendell, L.M. (2009). Chronic spinal hemisection in rats induces a progressive decline in transmission in uninjured fibers to motoneurons. Exp. Neurol. *216*, 471–480.

Center for Biologics Evaluation and Research. (2013). Guidance for Industry - Preclinical Assessment of Investigational Cellular and Gene Therapy Products (US Department of Health and Huma Services, US Food and Drug Administration and Center for Biologics Evaluation and Research, Office of Communication, Outreach, and Development (OCOD), US Government), p. 35.

Chang, J., Phelan, M., and Cummings, B.J. (2015). A meta-analysis of efficacy in pre-clinical human stem cell therapies for traumatic brain injury. Exp. Neurol. *273*, 225–233.

Chinnadurai, R., Ng, S., Velu, V., and Galipeau, J. (2015). Challenges in animal modelling of mesenchymal stromal cell therapy for inflammatory bowel disease. World J. Gastroenterol. *21*, 4779–4787.

Christopher and Dana Reeve Foundation (2008). One degree of separation. https://www.christopherreeve.org/living-with-paralysis/stats-about-paralysis.

Cummings, B.J., Uchida, N., Tamaki, S.J., Salazar, D.L., Hooshmand, M., Summers, R., Gage, F.H., and Anderson, A.J. (2005). Human neural stem cells differentiate and promote locomotor recovery in spinal cord-injured mice. Proc. Natl. Acad. Sci. USA *102*, 14069–14074.

Dressel, R., Schindehutte, J., Kuhlmann, T., Elsner, L., Novota, P., Baier, P.C., Schillert, A., Bickeboller, H., Herrmann, T., Trenkwalder, C., et al. (2008). The tumorigenicity of mouse embryonic stem cells and in vitro differentiated neuronal cells is controlled by the recipients' immune response. PLoS One *3*, e2622.

Fawcett, J.W., Curt, A., Steeves, J.D., Coleman, W.P., Tuszynski, M.H., Lammertse, D., Bartlett, P.F., Blight, A.R., Dietz, V., Ditunno, J., et al. (2007). Guidelines for the conduct of clinical trials for spinal cord injury as developed by the ICCP panel: spontaneous recovery after spinal cord injury and statistical power needed for therapeutic clinical trials. Spinal Cord *45*, 190–205.

Galipeau, J. (2013). The mesenchymal stromal cells dilemma–does a negative phase III trial of random donor mesenchymal stromal cells in steroid-resistant graft-versus-host disease represent a death knell or a bump in the road? Cytotherapy *15*, 2–8.

Gupta, N., Henry, R.G., Strober, J., Kang, S.M., Lim, D.A., Bucci, M., Caverzasi, E., Gaetano, L., Mandelli, M.L., Ryan, T., et al. (2012). Neural stem cell engraftment and myelination in the human brain. Sci. Transl. Med. *4*, 155ra137.

Hamers, F.P., Koopmans, G.C., and Joosten, E.A. (2006). CatWalk-assisted gait analysis in the assessment of spinal cord injury. J. Neurotrauma *23*, 537–548.

Henderson, V.C., Kimmelman, J., Fergusson, D., Grimshaw, J.M., and Hackam, D.G. (2013). Threats to validity in the design and conduct of preclinical efficacy studies: a systematic review of guidelines for in vivo animal experiments. PLoS Med. *10*, e1001489.

Hofstetter, C.P., Holmstrom, N.A., Lilja, J.A., Schweinhardt, P., Hao, J., Spenger, C., Wiesenfeld-Hallin, Z., Kurpad, S.N., Frisen, J., and Olson, L. (2005). Allodynia limits the usefulness of intraspinal neural stem cell grafts; directed differentiation improves outcome. Nat. Neurosci. *8*, 346–353.

Hooshmand, M.J., Sontag, C.J., Uchida, N., Tamaki, S., Anderson, A.J., and Cummings, B.J. (2009). Analysis of host-mediated repair mechanisms after human CNS-stem cell transplantation for spinal cord injury: correlation of engraftment with recovery. PLoS One *4*, e5871.

Houle, J.D., and Tessler, A. (2003). Repair of chronic spinal cord injury. Exp. Neurol. *182*, 247–260.

Hyun, I., Lindvall, O., Ahrlund-Richter, L., Cattaneo, E., Cavazzana-Calvo, M., Cossu, G., De Luca, M., Fox, I.J., Gerstle, C., Goldstein, R.A., et al. (2008). New ISSCR guidelines underscore major principles for responsible translational stem cell research. Cell Stem Cell *3*, 607–609.

Katz, R. (2004). FDA: evidentiary standards for drug development and approval. NeuroRx *1*, 307–316.

Kleinschnitz, C., Hofstetter, H.H., Meuth, S.G., Braeuninger, S., Sommer, C., and Stoll, G. (2006). T cell infiltration after chronic constriction injury of mouse sciatic nerve is associated with interleukin-17 expression. Exp. Neurol. *200*, 480–485.

Krassioukov, A., and Claydon, V.E. (2006). The clinical problems in cardiovascular control following spinal cord injury: an overview. Prog. Brain Res. *152*, 223–229.

Kwon, B.K., Soril, L.J., Bacon, M., Beattie, M.S., Blesch, A., Bresnahan, J.C., Bunge, M.B., Dunlop, S.A., Fehlings, M.G., Ferguson, A.R., et al. (2013). Demonstrating efficacy in preclinical studies of cellular therapies for spinal cord injury - how much is enough? Exp. Neurol. *248*, 30–44.

Lapchak, P.A. (2012). Scientific Rigor recommendations for optimizing the clinical applicability of translational research. J. Neurol. Neurophysiol. *3*, e111.

Lo, B., Kriegstein, A., and Grady, D. (2008). Clinical trials in stem cell transplantation: guidelines for scientific and ethical review. Clin. Trials *5*, 517–522.

Lucin, K.M., Sanders, V.M., Jones, T.B., Malarkey, W.B., and Popovich, P.G. (2007). Impaired antibody synthesis after spinal cord

injury is level dependent and is due to sympathetic nervous system dysregulation. Exp. Neurol. *207*, 75–84.

Macias, M.Y., Syring, M.B., Pizzi, M.A., Crowe, M.J., Alexanian, A.R., and Kurpad, S.N. (2006). Pain with no gain: allodynia following neural stem cell transplantation in spinal cord injury. Exp. Neurol. *201*, 335–348.

Marsh, S., Yeung, S., Torres, M., Lau, L., Davis, J., Capela, A., Monuki, E., Poon, W., and Blurton-Jones, M. (2017). Human neural stem cells fail to terminally differentiate, cause tumor-like growths, and provide no cognitive benefits in an immune-deficient transgenic model of Alzheimer's disease. Stem Cell Rep. *8*, this issue, 235–248.

Moalem, G., Xu, K., and Yu, L. (2004). T lymphocytes play a role in neuropathic pain following peripheral nerve injury in rats. Neuroscience *129*, 767–777.

Moller, D.R. (2014). Negative clinical trials in sarcoidosis: failed therapies or flawed study design? Eur. Respir. J. *44*, 1123–1126.

Nishi, R.A., Liu, H., Chu, Y., Hamamura, M., Su, M.Y., Nalcioglu, O., and Anderson, A.J. (2007). Behavioral, histological, and ex vivo magnetic resonance imaging assessment of graded contusion spinal cord injury in mice. J. Neurotrauma *24*, 674–689.

Pawar, K., Cummings, B.J., Thomas, A., Shea, L.D., Levine, A., Pfaff, S., and Anderson, A.J. (2015). Biomaterial bridges enable regeneration and re-entry of corticospinal tract axons into the caudal spinal cord after SCI: association with recovery of forelimb function. Biomaterials *65*, 1–12.

Perrin, S. (2014). Preclinical research: make mouse studies work. Nature *507*, 423–425.

Piltti, K.M., Salazar, D.L., Uchida, N., Cummings, B.J., and Anderson, A.J. (2013a). Safety of epicenter versus intact parenchyma as a transplantation site for human neural stem cells for spinal cord injury therapy. Stem Cells Transl. Med. *2*, 204–216.

Piltti, K.M., Salazar, D.L., Uchida, N., Cummings, B.J., and Anderson, A.J. (2013b). Safety of human neural stem cell transplantation in chronic spinal cord injury. Stem Cells Transl. Med. *2*, 961–974.

Piltti, K.M., Avakian, S.N., Funes, G.M., Hu, A., Uchida, N., Anderson, A.J., and Cummings, B.J. (2015). Transplantation dose alters the dynamics of human neural stem cell engraftment, proliferation and migration after spinal cord injury. Stem Cell Res. *15*, 341–353.

Roberts, S.F., Fischhoff, M.A., Sakowski, S.A., and Feldman, E.L. (2012). Perspective: transforming science into medicine: how clinician-scientists can build bridges across research's "valley of death". Acad. Med. *87*, 266–270.

Salazar, D.L., Uchida, N., Hamers, F.P., Cummings, B.J., and Anderson, A.J. (2010). Human neural stem cells differentiate and promote locomotor recovery in an early chronic spinal cord injury NOD-scid mouse model. PLoS One *5*, e12272.

Selden, N.R., Al-Uzri, A., Huhn, S.L., Koch, T.K., Sikora, D.M., Nguyen-Driver, M.D., Guillaume, D.J., Koh, J.L., Gultekin, S.H., Anderson, J.C., et al. (2013). Central nervous system stem cell transplantation for children with neuronal ceroid lipofuscinosis. J. Neurosurg. Pediatr. *11*, 643–652.

Sena, E.S., van der Worp, H.B., Bath, P.M., Howells, D.W., and Macleod, M.R. (2010). Publication bias in reports of animal stroke studies leads to major overstatement of efficacy. PLoS Biol. *8*, e1000344.

Sontag, C.J., Nguyen, H.X., Kamei, N., Uchida, N., Anderson, A.J., and Cummings, B.J. (2013). Immunosuppressants affect human neural stem cells in vitro but not in an in vivo model of spinal cord injury. Stem Cells Transl. Med. *2*, 731–744.

Sontag, C.J., Uchida, N., Cummings, B.J., and Anderson, A.J. (2014). Injury to the spinal cord niche alters the engraftment dynamics of human neural stem cells. Stem Cell Rep. *2*, 620–632.

Stokes, B.T., and Jakeman, L.B. (2002). Experimental modelling of human spinal cord injury: a model that crosses the species barrier and mimics the spectrum of human cytopathology. Spinal Cord *40*, 101–109.

van der Worp, H.B., Howells, D.W., Sena, E.S., Porritt, M.J., Rewell, S., O'Collins, V., and Macleod, M.R. (2010). Can animal models of disease reliably inform human studies? PLoS Med. *7*, e1000245.

Zhang, Y., Guan, Z., Reader, B., Shawler, T., Mandrekar-Colucci, S., Huang, K., Weil, Z., Bratasz, A., Wells, J., Powell, N.D., et al. (2013). Autonomic dysreflexia causes chronic immune suppression after spinal cord injury. J. Neurosci. *33*, 12970–12981.

Stem Cell Reports, Volume *8*

Supplemental Information

Preclinical Efficacy Failure of Human CNS-Derived Stem Cells for Use in the Pathway Study of Cervical Spinal Cord Injury

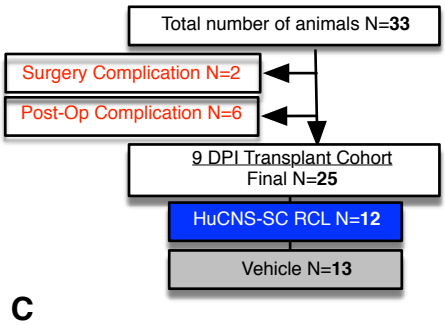
Aileen J. Anderson, Katja M. Piltti, Mitra J. Hooshmand, Rebecca A. Nishi, and Brian J. Cummings

Summary of CCL cell shipment from StemCells Inc					
Day	Date	TempTale temperature recordings			Cell Viability
		Temperature	SunShine	Alarm	
Day 1	3/6/2013	6˚C	Y	Y	94.6%
Day 2	3/7/2013	5˚C	Y	N	90.9%
Day 3	3/8/2013	5˚C	Y	N	90.6%
Day 4	3/20/2013	5˚C	Y	N	88.0%
Day 5	3/21/2013	5˚C	Y	N	86.0%
Day 6	3/22/2013	5˚C	Y	N	81.5%
Day 7	4/24/2013	5˚C	Y	N	89.3%
Day 8	4/25/2013	3˚C	Y	N	81.6%
Day 9	4/26/2013	5˚C	Y	N	89.8%
Day 10	5/8/2013	3˚C	Y	N	84.2%
Day 11	5/9/2013	3˚C	Y	N	87.9%
Day 12	5/10/2013	3˚C	Y	N	88.9%

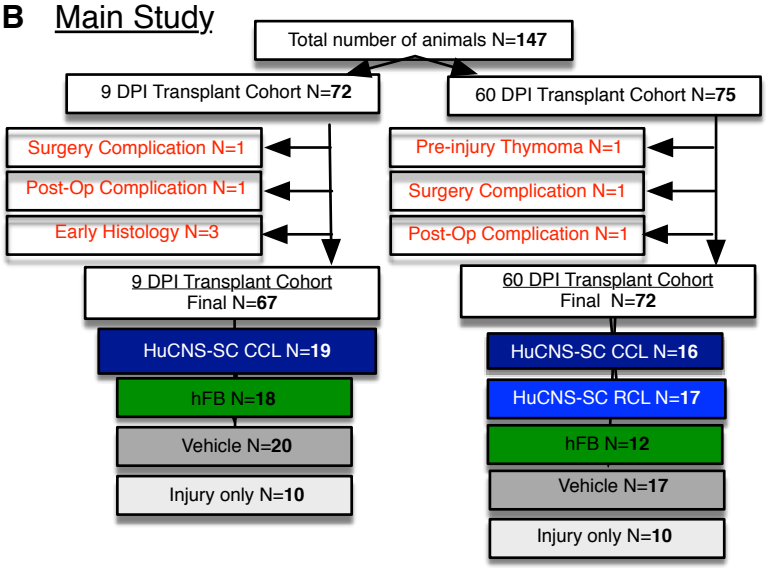
Supplemental Table S1. Summary of CCL cell shipments received from StemCells Inc. HuCNS-SC CCL overnight delivery included tracking of 24hr temperature maintenance and alarm conditions. 12 shipments (Day) were received over two months (Date). TempTale temperature recordings indicate the average temperature, Sunshine Y indicates temperature data was recorded continuously throughout the shipment, Alarm Y or N indicates whether temperature remained stable throughout the shipping. One CCL shipment exhibited a TempTale alarm Y upon arrival at UCI, as well as an extreme amount of cell debris, precluding transplantation. Transplantation surgeries were rescheduled and re-randomized to exclude using the Day 1 shipment. Cell viability was assessed each day prior to transplant via trypan blue exclusion.

Surgical and post-operative exclusions

A Proof-of-concept study - 9 DPI



B Main Study



C

Grubbs test exclusions and final animal numbers for statistical analysis

	Proof-of-concept study 9 DPI			
	HuCNS-SC RCL (N=12)*		Vehicle (N=13)*	
Histological analysis at 12 wpt	Excluded N	Final N	Excluded N	Final N
STEM121	1	5	-	-
Behavioral measure at 12 wpt	**Excluded N**	**Final N**	**Excluded N**	**Final N**
Ladder beam LF	0	12	1	12
Ladder beam RF	0	12	0	13

D

Grubbs test exclusions and final animal numbers for statistical analysis

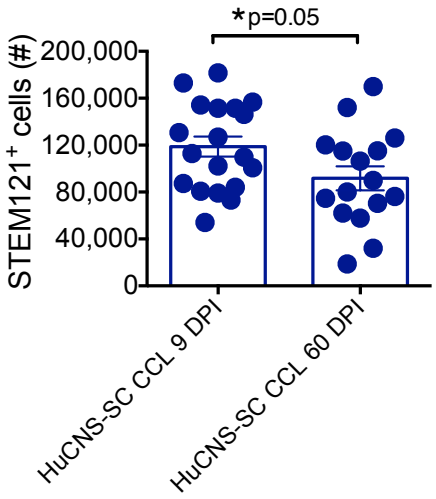
	9 DPI Transplant cohort								60 DPI Transplant cohort							
	HuCNS-SC CCL (N=19)*		hFB (N=18)*		Vehicle (N=20)*		Injured only (N=10)*		HuCNS-SC CCL (N=16)*		HuCNS-SC RCL (N=17)*		hFB (N=12)*		Vehicle (N=17)*		Injured only (N=10)*	
Histological analysis at 12 wpt	Excluded N	Final N	Excluded N	Final N	Excluded N	Final N	Excluded N	Final N	Excluded N	Final N	Excluded N	Final N	Excluded N	Final N	Excluded N	Final N	Excluded N	Final N
STEM121	0	19	0	18	-	-	-	-	0	16	0	17	0	12	-	-	-	-
STEM121/DCX	0	7	-	-	-	-	-	-	0	7	0	7	-	-	-	-	-	-
STEM123	0	7	-	-	-	-	-	-	1	6	0	7	-	-	-	-	-	-
STEM121/Olig2	0	7	-	-	-	-	-	-	1	5	0	7	-	-	-	-	-	-
STEM121/CC1	0	7	-	-	-	-	-	-	0	7	0	7	-	-	-	-	-	-
Behavioral measure at 12 wpt	**Excluded N**	**Final N**	**Excluded N**	**Final N**	**Excluded N**	**Final N**	**Excluded N**	**Final N**	**Excluded N**	**Final N**	**Excluded N**	**Final N**	**Excluded N**	**Final N**	**Excluded N**	**Final N**	**Excluded N**	**Final N**
Grip Strength LF	0	19	0	18	0	20	0	10	0	16	0	17	0	12	0	17	0	10
Grip Strength RF	0	19	0	18	0	20	0	10	1	15	0	17	0	12	0	17	0	10
Cylinder LF	0	19	0	18	1	19	0	10	1	15	0	17	0	12	0	17	0	10
Cylinder RF	1	18	1	17	0	20	0	10	1	15	1	16	1	11	1	16	0	10
Both Cylinder	0	19	1	17	1	19	0	10	1	15	0	17	0	12	0	17	0	10
Ladder beam LF	0	19	1	17	0	20	0	10	0	16	0	17	1	11	1	16	0	10
Ladder beam RF	1	18	1	17	1	19	1	9	1	15	0	17	1	11	0	17	0	10
Von Frey LF	0	19	0	18	0	20	0	10	0	16	0	17	0	12	0	17	0	10
Von Frey RF	0	19	1	17	0	20	0	10	0	16	0	17	0	12	0	17	0	10
Hargreaves LF	1	18	1	17	0	20	0	10	0	16	0	17	0	12	0	17	0	10
Hargreaves RF	0	19	0	18	1	19	0	10	0	16	0	17	0	12	0	17	0	10
Catwalk Run Duration	0	19	0	18	1	19	0	10	0	16	0	17	0	12	1	16	0	10
Catwalk swing speed RF	0	19	0	18	1	19	0	10	1	15	0	17	0	12	1	16	0	10
Catwalk swing speed LF	1	18	1	17	1	19	0	10	0	16	0	17	0	12	1	16	0	10
Catwalk duty cycle RF	0	19	0	18	1	19	0	10	0	16	0	17	0	12	2	15	0	10
Catwalk duty cycle LF	0	19	0	18	1	19	0	10	0	16	0	17	0	12	1	16	0	10
Catwalk BOS F	1	18	0	18	1	19	0	10	0	16	0	17	1	11	1	16	0	10
Catwalk RI	0	19	1	17	1	19	0	10	1	15	0	17	1	11	1	16	0	10
Catwalk Aa	1	18	1	17	2	18	1	9	1	15	1	16	1	11	1	16	0	10
Catwalk Ab	0	19	0	18	1	19	0	10	0	16	0	17	0	12	1	16	0	10
Catwalk Ca	0	19	0	18	1	19	0	10	0	16	0	17	0	12	1	16	0	10
Catwalk Cb	1	18	1	17	2	18	0	10	1	15	1	16	0	12	2	15	1	9
Catwalk Ra	1	18	1	17	2	18	0	10	0	16	1	16	1	11	2	15	0	10
Catwalk Rb	0	19	1	17	2	18	0	10	1	15	0	17	1	11	1	16	1	9

Supplemental Figure S1. Surgical and post-operative exclusions, Grubb's outlier exclusions, and final group numbers. A-B) Surgical and post-operative exclusions for the proof-of-concept experiment and the main study. C) Proof-of concept study Grubb's outlier exclusions and final N for each statistical analysis. STEM121+ donor human cell engraftment was assessed in 7 randomly selected animals (Methods). D) Main study Grubb's outlier exclusions and final N for each statistical analysis. Donor human cell fate analysis was performed in 7 randomly chosen animals (Methods). N/A indicates not applicable, as groups had either no surviving cells or no transplants.

Antibody	Dilution	Host	Company	Catalog #
SC121	1:10,000	Mouse	Stem Cells Inc	AB121U059
SC123	1:2,500	Mouse	Stem Cells Inc	AB123U050
APC/CC1	1:200	Mouse	CalbioChem	OP80
DCX	1:50	Goat	Santa Cruz Biotech	SC-8066
Olig2	1:500	Goat	R&D Systems	AF2418
Anti-Mouse F(ab')2	1:500	Donkey	Jackson ImmunoResearch	715066151
Anti-Goat F(ab')2	1:500	Donkey	Jackson ImmunoResearch	705066147

Supplemental Table S2. List of Antibodies. Antibodies, dilutions, host, source, and catalogue number for reagents used in histological analyses. Nuclear counterstaining with Hematoxylin or Methyl green as indicated in figures.

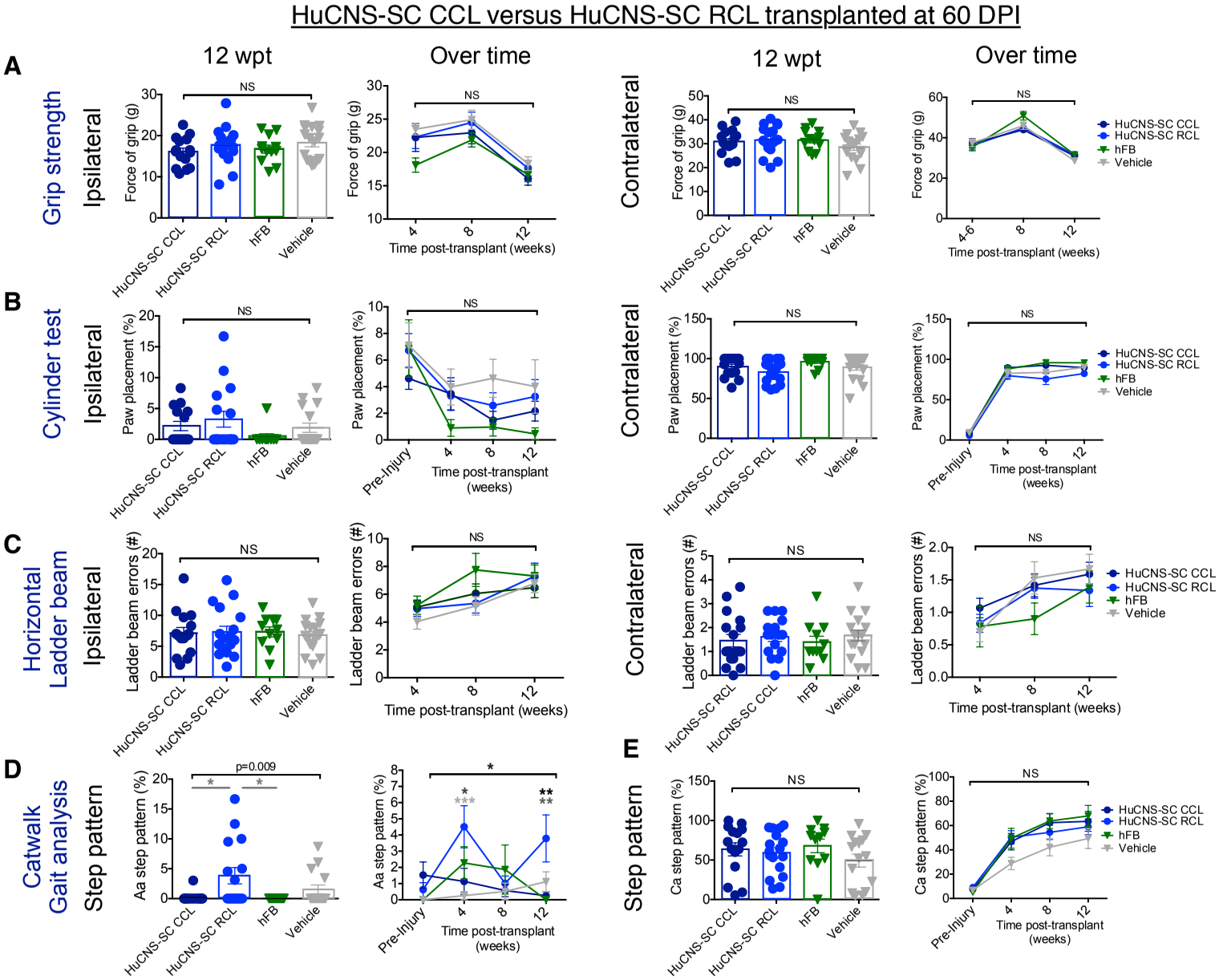
A



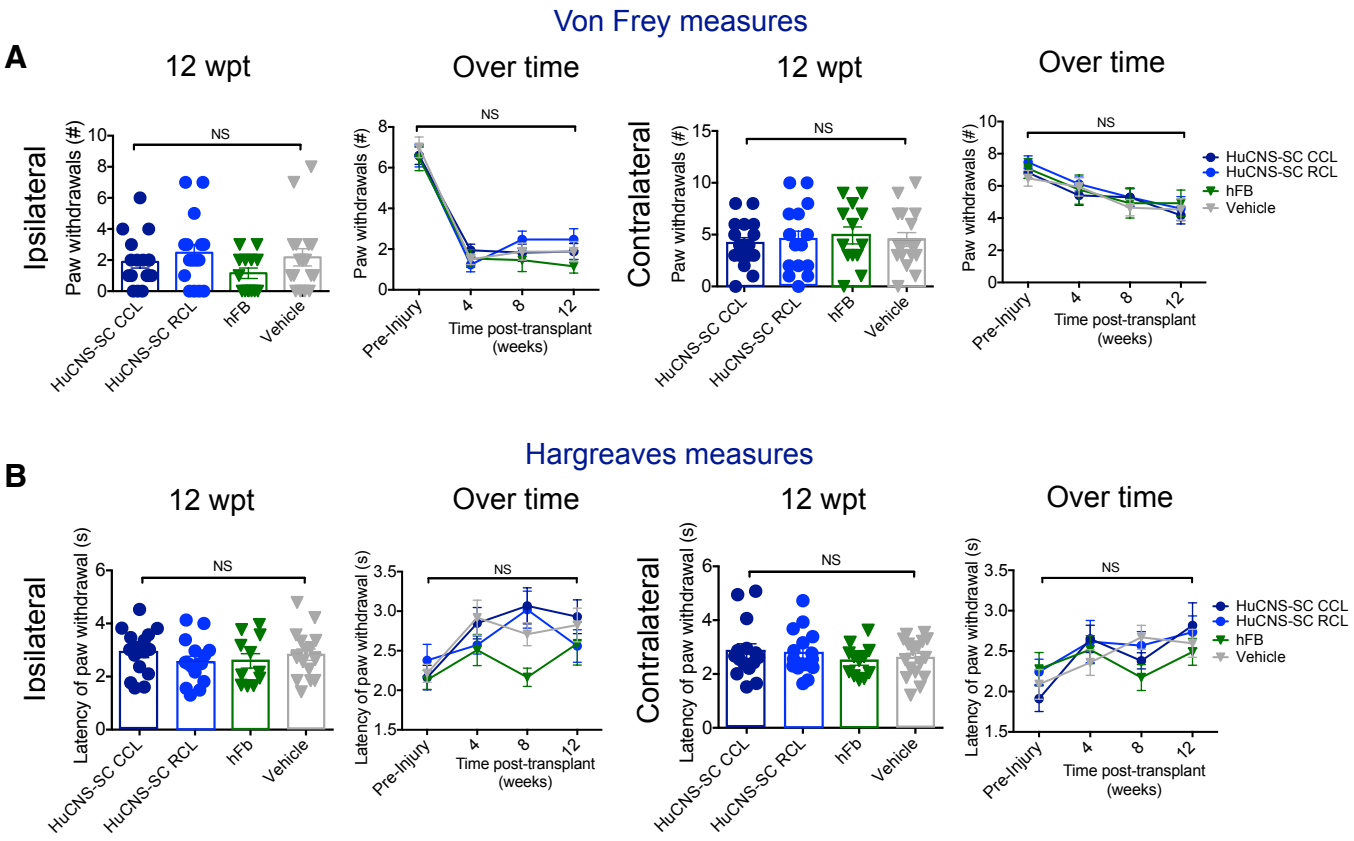
B

Staining	Lineage	Transplantation groups: HuCNS-SC CCL 9 DPI	Transplantation groups: HuCNS-SC CCL 60 DPI	Student's 2-tailed t-test p-value
STEM121/DCX	Early neuronal	0.0% ± 0.0%	0.0% ± 0.0%	NS
STEM123	Astroglial	33.0% ± 4.4%	16.3% ± 2.4%	p=0.009**
STEM121/Olig2	Oligodendroglial	55.6% ± 1.6%	54.2% ± 0.8%	p=0.2
STEM121/CC1	Mature Oligodendroglial	10.1% ± 2.0%	13.6% ± 3.0%	p=0.4

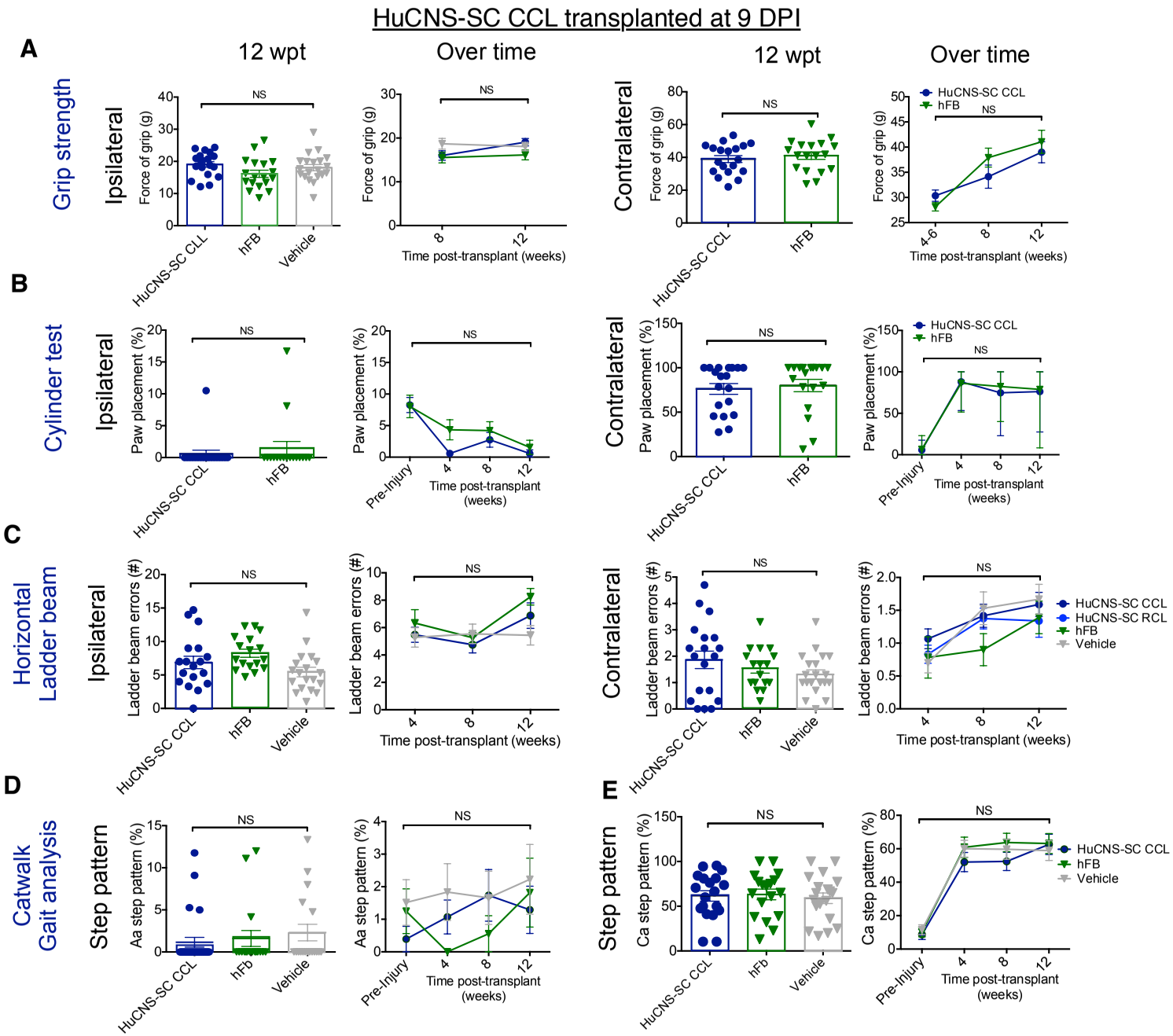
Supplemental Figure S2. Comparison of donor human cell engraftment and lineage in 60 DPI versus 9 DPI HuCNS-SC CCL transplantation cohort groups. A) STEM121+ donor human cell engraftment was significantly increased in the HuCNS-SC CCL 9 DPI (n=19) versus 60 DPI (n=16) cohorts at 12 WPT (Student's two-tailed t-test, $p<0.05$). B) Comparison of donor human cell lineage proportions between the HuCNS-SC CCL 60 DPI and 9 DPI cohort groups revealed no significant differences in early neuronal cells or oligodendroglial cells. However, the proportion of STEM123+ astroglial marker positive cells was significantly increased in the HuCNS-SC CCL 9 DPI compared to the 60 DPI cohort group (Student's two-tailed t-test, $p<0.009$). N=5-7 for 60 DPI lineage vs n=7 for 9 DPI lineage analysis (see Supplemental Figure S1.D for specific numbers).



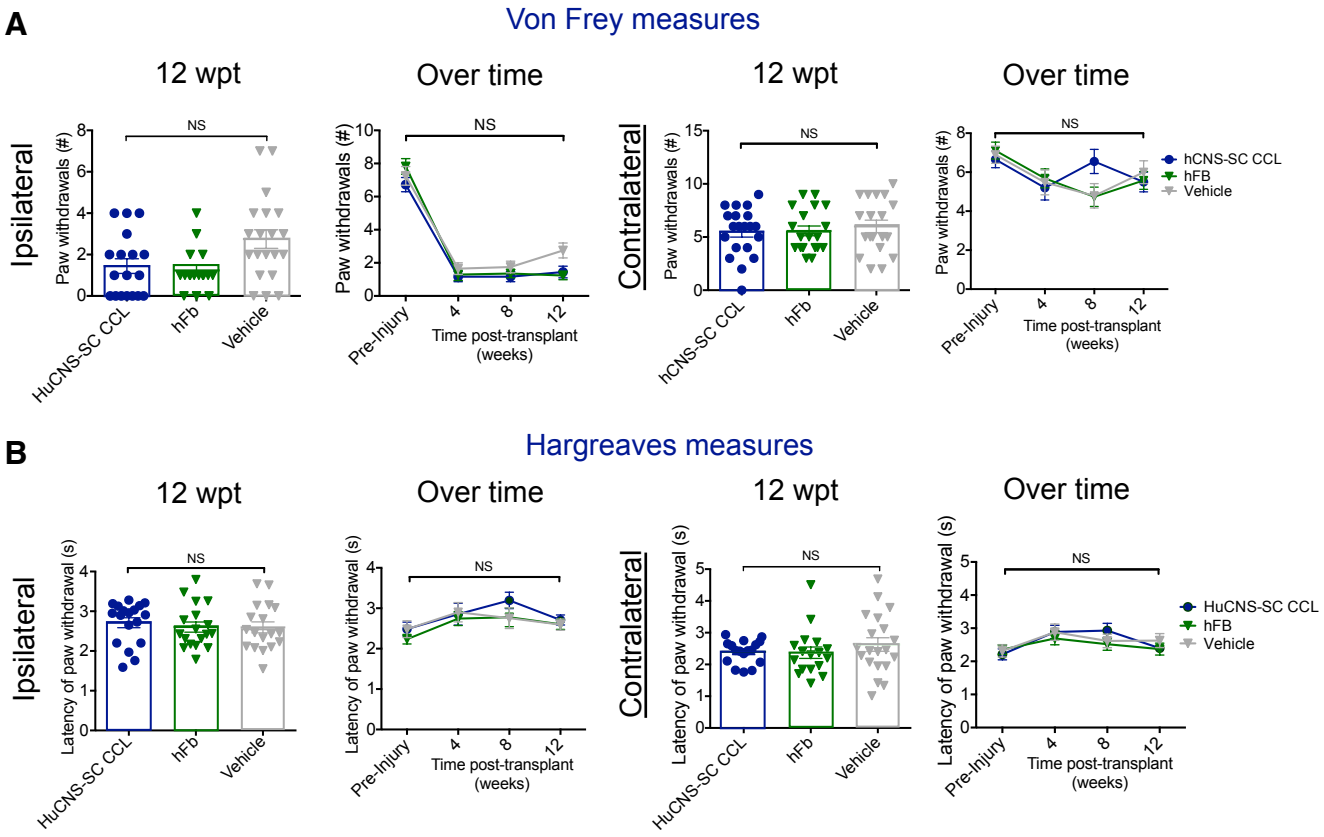
Supplemental Figure S3. Comparison of locomotor recovery in HuCNS-SC CCL and RCL 60 DPI transplantation cohorts at all assessment times. No differences between groups were found in **A)** ipsilateral or contralateral forelimb grip strength, **B)** ipsilateral or contralateral forelimb cylinder reaching (paw placement %), **C)** ipsilateral or contralateral forelimb horizontal ladder beam errors, or **D)** Catwalk Aa step pattern was significantly increased (1-way ANOVA, $p<0.009$) in the HuCNS-SC RCL group in comparison with the CCL and hFB groups (Tukey's t-test, $p<0.05$) at 12 WPT. In parallel, Aa step pattern was changed across time (2-way ANOVA interaction effect, black bracket; *$p\leq0.05$), and post-hoc Bonferroni t-tests revealed significantly increased Aa step pattern for the HuCNS-SC RCL group compared to the vehicle (light grey asterisks, *$p\leq0.05$) and CCL groups (black asterisks, **≤ 0.001) at 4 WPT, and compared to the hFB (grey asterisks, **$p\leq0.01$) and CCL groups (black asterisks, *≤ 0.05) at 12 WPT. **E)** Catwalk Ca step pattern (the predominant post-injury step pattern) was not significantly changed at either 12 WPT (1-way ANOVA, NS $p>0.05$) or at any other assessment time (2-way ANOVA, NS $p>0.05$). *Grip*, n=16, 17, 12, 17; *Cylinder*, n=15, 17, 12, 17; *Ladder*, n=16, 17, 11, 16; and *Catwalk*, n=15, 16, 11, 16 for CCL, RCL, hFB and Vehicle respectively (see Supplemental Figure S1.D).



Supplemental Figure S4. Comparison of sensory parameters for ipsilateral and contralateral forelimbs within HuCNS-SC CCL and RCL 60 DPI transplantation cohort groups at all assessment times. A) No differences were found between groups in ipsilateral or contralateral forelimb withdrawal number in Von Frey testing, or **B)** withdrawal latency in Hargreaves testing at either 12 WPT (1-way ANOVA, NS $p>0.05$) or at any other assessment time (2-way ANOVA, NS $p>0.05$). *Von Frey*, n=16, 17, 12, 17; *Hargreaves*, n=16, 17, 12, 17; for CCL, RCL, hFB and Vehicle respectively (see Supplemental Figure S1.D).



Supplemental Figure S5. Comparison of locomotor recovery within HuCNS-SC CCL 9 DPI transplantation cohort groups at all assessment times. No differences were found in A) ipsilateral or contralateral forelimb grip strength, B) ipsilateral or contralateral forelimb cylinder reaching (paw placement %), C) ipsilateral or contralateral forelimb horizontal ladder beam errors, D) Catwalk Aa step pattern, or E) Ca step pattern between the study groups at 12 wpt (1-way ANOVA, $p>0.05$) or at any other assessment time (2-way ANOVA, $p>0.05$). *Grip*, n=19, 18, 20; *Cylinder*, n=19, 18, 19; *Ladder*, n=19, 17, 20; and *Catwalk*, n=18, 17, 18, for CCL, hFB, and Vehicle respectively (see Supplemental Figure S1.D).



Supplemental Figure S6. Comparison of sensory parameters for ipsilateral and contralateral forelimbs within HuCNS-SC CCL 9 DPI transplantation cohort groups at all assessment times. A) No differences were found between groups in ipsilateral or contralateral forelimb withdrawal number in Von Frey testing at 12 WPT or any timepoint, or **B)** withdrawal latency in Hargreaves testing at 12 WPT (1-way ANOVA, NS $p>0.05$) or at any timepoint (2-way ANOVA, NS $p>0.05$). *Von Frey*, n=19, 18, 20; *Hargreaves*, n=18, 17, 20; for CCL, hFB and Vehicle respectively (see Supplemental Figure S1.D).

Supplemental Experimental Procedures

HuCNS-SC and hFb

HuCNS-SC are derived via fluorescence-activated cell sorting (FACS) from donated fetal brain tissue and expanded as neurospheres based on expression of the stem cell marker CD133+, a lack of the hematopoietic markers; CD34- and CD45-, and low levels of CD24lo. Sorting by these markers has previously been shown to result in a highly enriched population of human neural stem cells (Uchida et al. 2000). Sorted HuCNS-SC CCL and RCL cells were provided by StemCells Inc (StemCells Inc., Palo Alto, CA, http://www.stemcellsinc.com).

HuCNS-SC RCL cells were maintained and shipped to UCI via overnight delivery from the research division of StemCells Inc. in a volume of 50ml. After receipt of cells on the day of transplantation, neurospheres (< passage number 12 on arrival) were dissociated into individual cells, centrifuged, washed, counted under sterile conditions, and adjusted to a cell density of 75,000 cells per microliter in X-Vivo 15 medium (Lonza) for injection, as previously described (Cummings et al. 2005, Cummings et al. 2006, Hooshmand et al. 2009, Salazar et al. 2010, Piltti et al. 2013, Piltti et al. 2013, Sontag et al. 2013, Sontag et al. 2014, Piltti et al. 2015). Each day of surgery was performed with a new RCL shipment. Cell yield, viability, and preparation data were recorded for each vial of cells received.

HuCNS-SC CCL seed stock were produced and maintained by StemCells Inc. under cGMP/GTP protocols and conditions. CCL cells sent to UCI for transplantation were prepared in the StemCells Inc. process development laboratory in a non-cGMP environment. Cell passage numbers were not provided by StemCells Inc. for CCL cells received at UCI. HuCNS-SC CCL cells were shipped to UCI via overnight delivery per an established Technical Research and Development Protocol with StemCells Inc. Specifically, each cell shipment was sent in a volume of 300µl, monitored using TempTale, TiltWatch Plus, and ShockWatch devices, and all data from shipment monitoring recorded. HuCNS-SC CCL cells received at UCI were designated as 'non-clinical product' cells on the vials received. After receipt of cells on the day of transplantation, neurospheres were dissociated into individual cells and adjusted to a cell density of 75,000 cells per microliter in X-Vivo 15 medium (Lonza) for injection, as detailed for RCL cells. One CCL shipment exhibited a TempTale alert upon arrival at UCI as well as extensive cellular debris, precluding transplantation; transplantation surgeries were rescheduled to exclude cells from that day's shipment. Each day of surgery was performed with a new CCL shipment, thus, 12 CCL shipments were received, and 1 was excluded as noted. Cell yield, beginning and end of day viability, and cell preparation data were recorded for each vial of cells received (Supplemental Table 1).

Human mesenchymal stromal cell-fibroblasts (hFb) (Cell Applications, San Diego, CA, http://www.cellapplications.com) were thawed and cultured in DMEM supplemented with 10% fetal bovine serum and Glutamine at UCI for 7 days prior to transplantation. For transplantation, hFb at passage number 8 were dissociated into individual cells and adjusted to a cell density of 75,000 cells per microliter in X-Vivo 15 medium (Lonza).

Animal Welfare

This study was carried out in accordance with the Institutional Animal Care and Use Committee at the University of California, Irvine, and was consistent with current U.S. federal guidelines.

Contusion Injuries

For the proof-of-concept study cohort, commercially available Rag2γ(c) female mice (Taconic Biosciences, http://www.taconic.com) were used at 18 months of age. For the main study cohorts, adult female 10-12 week-old Agouti Rag2γ(c) hybrid mice, generated at StemCells Inc were used. For all cohorts, animals were anesthetized with 2% isoflurane (VetEquip Inc., Pleasanton, CA, http://www.vetequip.com). Spinal cords at cervical 5 (C5) vertebral level were exposed by laminectomy using a surgical microscope and stabilized in a spinal stereotaxic frame by clamping at the C4 and C6 lateral vertebral processes, and unilateral 30-kDa contusion injuries with 5s dwell time were administrated with a 1mm diameter impactor tip on right side of the spinal cord, between the midline and lateral edge of the C5 vertebrae using an Infinite Horizon Impactor (Precision Systems and Instrumentation, Lexington, KY, http://www.presysin.com) as previously described (Nishi et al. In review). Following the injury, the exposed spinal cords were covered with gelfoam (Pfizer, New York, NY, http://www.pfizer.com), muscles were closed with 5-0 chromic gut sutures (Surgical Specialties Co., Reading, PA, http://www.heidolphna.com), and the skin was closed using wound clips (CellPoint Scientific Inc., Gaithersburg, MD, http://www.cellpointscientific.com). For postoperative care, the animals received 0.01 mg/kg s.c buprenorphine (Hospira Inc., Lake Forest, IL, http://www.hospira.com) twice a day for 2 days, 50 ml/kg SQ lactated Ringer's solution (B. Braun Medical Inc., Irvine, CA, http://www.bbraunusa.com) once daily for 4 days, and manual bladder expression twice a day until mice recovered some level of bladder function, then once daily until the end of study. All the animals were maintained on antibiotics rotating 2.5mg/kg SQ Enrofloxacin (Baytril) (Western Medical), 2.5mg/kg p.o. Ciprofloxacin hydrochloride (Dr. Reddy's Laboratories, Bachepalli, India, http://www.drreddys.com), and 2.5mg/kg p.o. Ampicillin (STADA Pharmaceuticals, Cranbury, NJ, http://www.stada.de/english) every 2 weeks until end of the study.

HuCNS-SC and hFibroblast Transplantation

Mice were re-anesthetized 9 or 60 DPI, laminectomy sites re-exposed, and vertebral column stabilized in a spinal stereotaxic frame for injection. A total volume of 1µl (250nl per injection site) of cell suspension or vehicle (X-Vivo 15 medium) was injected in two rostral bilateral injections and another two bilateral caudal injections, 0.75 mm from midline, 1mm distal (rostral or caudal respectively) to the injury site, using polished 30° beveled glass pipettes (inner diameter [i.d.] 70µm, outer diameter [o.d.] 100–110µm; Sutter Instruments, Novato, CA, http://www.sutter.com). Injections were performed using a NanoInjector 2000 system with a Micro4 Controller and a micropositioner (World Precision Instruments, Waltham, MA, http://www.wpiinc.com), under microscopic guidance over 1 minute, followed by an additional 2-minute delay before removing the needle to prevent back-flow. After injection, the postoperative procedures and animal cal were performed as described.

Data management, Exclusions, Final Ns, and experimental blinding.

All data for cell shipments, animal surgeries, pre- and post-operative care, behavioral assessments, and perfusions / staining were maintained under pre-established GLP-like protocols, with an assigned data monitor. All injuries and transplantations were done by well-trained personnel. Consistency of the injuries was validated by IH device feedback (actual force and displacement values over time) followed by behavioral outcome monitoring. All animal care, behavioral data collection and analysis was performed by investigators blinded to the study groups. To

maintain blinding, animals were randomized for distribution into groups by one investigator the night before transplantation, a second investigator was responsible for animal coding and distribution of vehicle, hFB, or HuCNS-SC aliquots for injection based on the pre-established group assignments. Only these two investigators had access to the code. All investigators conducting anesthesia, surgery, transplantation, behavior, and histology remained blinded for the duration of the study.

Criteria for inclusion/exclusion were established prior to conducting either study. 33 animals entered the proof of concept study cohort; 25 completed the study. 147 animals entered the main study cohorts; 139 completed the study. There were no animal exclusions due to engraftment failure or at the stage of histological analysis. Pre- and post-injury animal exclusions, Grubbs test exclusions, and final animal numbers for statistical analysis are detailed in *Supplemental Fig 1A-D*.

Assessment of locomotor and sensory function

Horizontal Ladder Beam task was performed as previously described (Cummings et al. 2007) at 8, and 12 weeks post-transplantation (WPT). Briefly, the number of contralateral (left) and ipsilateral (right) forelimb stepping errors were analyzed in three separate runs per mouse across a horizontal ladder with 50 rungs. Stepping errors include missing the rung, stepping on the run with the dorsal surface of the paw, or slipping off of the rung after placing with the plantar surface of the paw. Successful steps include stepping squarely on the rung with the plantar surface of the paw, and skipping over the rung. Catwalk Gait (Vrinten and Hamers 2003, Hamers et al. 2006) was assessed in three separate runs per mouse as previously described (Salazar et al. 2010) using CatWalk XT (Noldus v9.0) prior to injury (baseline) and at 4, 8, and 12 WPT.

Forelimb-use asymmetry was assessed using Cylinder task as previously described (Khaing et al. 2012, Pawar et al. 2015) prior to injury (baseline) and at 4, 8, and 12 WPT. Briefly, mice were placed in a glass beaker and the number of forepaw placements (single and both paws) on the sides of beaker were counted over 5 minutes. Forepaw grip strength was measured for each paw alone and together in five trials per mouse using a Dunnett-style grip strength meter (Dunnett et al. 1998) as previously described (Pawar et al. 2015) at 4, 8, and 12 WPT.

Mechanical allodynia was assessed using Von Frey test as previously described (Salazar et al. 2010). Briefly, mice were placed in a clear acrylic chamber on an elevated wire mesh grid. Withdrawal response of the hindpaws was assessed by applying a 4.08 gram force Touch-Test Sensory Evaluator filament (North Coast Medical, Gilroy, CA, https://www.ncmedical.com) prior to injury (baseline) and at 4, 8, and 12 WPT. Filaments were administered to the plantar surface of each paw 10 times, 2 minutes apart, and the number of withdrawals was recorded. Thermal hyperalgesia was assessed using Hargreaves test (Hargreaves, Pain, 1988) as previously described (Piltti et al. 2013). Briefly, forepaw sensitivity was tested while mice were standing on top of a temperature-controlled glass plate heated to 35°C. A withdrawal response of all four paws were assessed using a radiant thermal stimulus of the paw analgesia meter set at an active intensity of 35 arbitrary units applied to the plantar surface through the glass plate (IITC Life Sciences, Inc, Woodland Hills, CA, http://www.iitcinc.com) prior to injury (baseline) and at 4, 8, and 12 WPT. Thermal stimulus was administered to plantar surface of each paw three times, with a 3 minute rest between each run, and the reaction times were recorded and then averaged. For both Von Frey and Hargreaves, animals were acclimatized to the testing chambers for1.5 to 2 hours prior to testing.

Perfusion, Tissue Collection, Sectioning and Histology

At 12 WPT, mice were terminally anesthetized and transcardially perfused, injured cord segments were dissected based on dorsal spinal root counts (C1-T2 roots), postfixed and cryoprotected with 20% sucrose, flash frozen in cooled isopentane, and stored for sectioning as previously described (Hooshmand et al. 2009). Injured cord segments from all animals were dissected and coronal sections of 30µm were taken using a cryostat (Thermo Fisher Scientific, Waltham, MA, http://www.thermofisher.com) followed by mounting onto slides using a CryoJane tape transfer system (Leica Biosystems, Inc., Buffalo Grove, IL, http://leicabiosystems.com). For immunohistochemistry, the tissue sections were antigen-retrieved in Buffer A (Electron Microscopy Sciences, Hatfield, PA http://www.emsdiasum.com) using 2100 Retriever (Aptum Biologics, South Hampton, http://www.aptum-bio.com/United Kingdom) and treated to deactivate endogenous peroxide activity, immunostained as previously described (Hooshmand et al. 2009) and visualized using either 3,3-diaminobenzidine (DAB) horseradish peroxidase or/and SG horseradish peroxidase substrates (Vector Laboratories, Burlingame, CA, http://vectorlabs.com). The primary antibodies and secondary antibodies used are listed in *Supplemental Table 2*. Methyl green or hematoxylin were used for nuclear labeling where necessary.

Stereological Quantification

Final animal numbers for analysis are as described above and listed in *Supplemental Fig 1*. Total numbers of STEM121$^+$ (SC121) human cells in all transplanted animals and STEM123$^+$ (SC123) human GFAP$^+$ cells were determined by unbiased stereology in 1 in 12 intervals from spinal cord sections 360 µm apart using systematic random sampling with an optical fractionator probe and StereoInvestigator version 11 (MicroBrightField Inc., Williston, VT, http://www.mbfbioscience.com). Optical fractionator grid size and counting frame size were empirically determined to yield average Gundersen (m=1) cumulative error values less than 0.1. The migration of human cells was analyzed as percentage of the cells per section relative to total number of counted STEM121$^+$ human cells. The distribution of the cells was normalized with the distance from the injury epicenter, designated as the most damaged tissue section with largest injury epicenter.

A random subset of animals were selected for proportional counting of cell fate (N=7/group). Proportional counts of STEM121$^+$/DCX$^+$, STEM121$^+$/Olig2$^+$, and STEM121$^+$/APC(CC-1)$^+$ cells were analyzed in 1 in 24 intervals from spinal cord sections 720 µm apart using an optical fractionator probe and systematic random sampling to accumulate a minimum of 100 targets; proportions were thus computed for the second label in reference to the STEM121$^+$ labeling in the same sections. STEM123$^+$ counts of human GFAP$^+$ cell fate were proportionally compared to STEM121$^+$ staining within separate sections /set of the same animal.

Statistical Analysis

All data are shown as mean±SEM; statistical analysis was performed using Prism software, version 6 (GraphPad Software Inc., San Diego, CA, http://www.graphpad.com). Comparisons between groups were analyzed using either one-way ANOVA combined with Tukey's post hoc t-tests or Student's one or two--tailed t-tests. In the case of behavioral data, where there was an *a priori* prediction that human cell groups would perform better than injured controls or vehicle groups, a one-tailed test was applied, as indicated in the text and legends. In the case of histological analyses, where there were no *a priori* predictions for comparisons, a two-tailed test was applied, as

indicated in the text and legends. Changes in locomotor and sensory function between the groups were compared using two-way repeated measures analysis of variance (ANOVA) combined with multiple comparison corrected/ Bonferroni post-hoc t-tests. Differences in migration were analyzed using unpaired two-tailed t-tests with Holm-Sidak multiple comparison correction as there was no *a priori* hypothesis at one cell type would behave differently than another. Correlation between numbers of $STEM121^+$ or $STEM121^+/CC1^+$ cells and numbers of Ladder beam errors or percentage of Ab or Ca step patterns were assessed using the Pearson correlation coefficient. A p value of ≤ 0.05 was considered to be statistically significant.

Power analyses were conducted prior to initiating the CCL cohorts. Power analysis based on preliminary efficacy data with the RCL indicated that a sample size of 10 in each group has 80% power at an alpha < 0.05 (two-tailed) to detect a change of 2% in CatWalk Aa step pattern, a 1 error reduction in contralateral horizontal ladder errors, and a 4 error reduction in ipsilateral horizontal ladder errors. For sensory testing, power analysis was based on the published literature (Hofstetter et al. 2005), and indicated that a sample size of 16 in each group has 80% power at an alpha < 0.05 (two-tailed) to detect a change of 5g in withdrawal threshold in Von Frey testing, and a 0.4s change in withdrawal latency in Hargreaves testing. These represent detection of very small effect sizes with a high degree of sensitivity.