Supplementary Discussion

Our results need to be discussed within the framework of the entire field of cardiac regeneration, which is fairly contentious and discordant at the present day. Regardless of the source of cells (c-kit$^+$ cells, other progenitor cells or cardiomyocytes themselves), the most conservative estimate is that only 0.0083% of cardiomyocytes in the MI border zone of the mouse heart undergo DNA synthesis, suggestive of cardiomyocyte proliferation that would be part of any regenerative process [25]. Traditional labeling approaches should equally mark new cardiomyocyte formation from existing cardiomyocytes as well as from progenitor cell activity, assuming that the labeling period is long enough to show cardiomyocyte differentiation markers and still retain the DNA tagging event, or that newly formed cardiomyocytes from progenitor cells undergo at least 1 round of proliferation. In adult cats, isoproterenol was infused to cause cardiac injury for 10 days with 7 days of BrdU labeling. This analysis resulted in 0.02 to 0.08% of cardiomyocytes being labeled for DNA synthesis in the left ventricle, again suggesting that new cardiomyocyte formation is detectable but rare even after injury, with the caveat that BrdU labeling might only reflect DNA synthesis and not true cellular proliferation [26]. While higher levels of cardiomyocyte cell cycle activity have been reported in the injured mammalian heart [27], the overall consensus is that such events do happen and that the heart does contain a finite capacity for regeneration with new cardiomyocyte formation, especially after injury, albeit at low levels.

A study based on the worldwide pulse of ^{14}C labeling with the atomic bomb testing era showed that new cardiomyocytes can be formed in the normal unstressed human heart over time (with aging), with a baseline level of 0.4 – 1% per year [28]. Using an $^{15}N^+$ isotope labeling approach in the mouse, a baseline comparable level of 0.76% per year was calculated for cardiomyocyte renewal in the heart, which was significantly higher after injury [29]. However, much higher levels of cardiomyocyte turnover have been reported using the same ^{14}C worldwide pulse-labeling phenomenon, with values of approximately 7% of cardiomyocyte renewal per year in 20 to 40 year old individuals, and even higher levels again when heart disease was present [30]. Regardless of the exact quantification, all studies support the emerging consensus that the heart can indeed regenerate at a finite level, especially after an injury event. A critical assessment of the data we presented would suggest that endogenous c-kit$^+$ lineage cells are not the primary mechanism behind this already low level of renewal.

Given the persistence of discordant reports in the cardiac regeneration / stem cell field, we sought independent verification of our basic finding that new cardiomyocyte formation from endogenous c-kit$^+$ lineage cells is an extremely rare event. Hence we sent an array of unprocessed histological samples, in a blinded format, to an entirely independent laboratory with known expertise in this area, but with whom we have never previously interacted (*Dr Eduardo Marbán, listed as an author)*. Their findings basically mirrored our results and showed that while examples of eGFP$^+$ cardiomyocytes could be found in the hearts of adult Kit$^{+/MCM}$ x R-GFP mice (8 weeks of tamoxifen), it was a very rare event (Extended Data Fig. 8a, b, c).

One potential issue with the Kit$^{+/MCM}$ or Kit$^{+/Cre}$ lineage tracing approach is how it might compare with past antibody-dependent approaches for identifying c-kit$^+$ cells. In our hands the c-kit antibody identified very few cells in the adult heart, and these cells were always small and mononuclear with coincident real-time eGFPnls expression from the *Kit*-Cre allele (the recombination-dependent R-GFP allele was not present). These results together with antibody-based detection of c-kit expressing cells by immunohistochemistry, demonstrates that the *Kit* knock-in approach only marks currently or past c-kit expressing cells. However, the Cre-dependent labeling approach was not 100% efficient, with ≈80-90% labeling directed by the *Kit*-Cre allele and 60-70% efficiency directed by the *Kit*-MerCreMer allele in the presence of tamoxifen, which would slightly under-represent the number of cardiomyocytes that could be derived from c-kit$^+$ lineage cells.

The analysis of real time eGFPnls expression in the hearts of $Kit^{+/Cre}$ mice (no R-GFP reporter) showed no cardiomyocytes or endothelial cells at P0 or 4 weeks of age, indicating that these differentiated cells do not activate the *Kit* allele "ectopically" such as due to a chromosomal rearrangement or some sort of transcriptional mis-coding in rare cardiomyocytes. However, it is possible that the *Kit* allele could be transiently expressed for a short window of time during development in rare populations of cardiomyocytes scattered throughout the heart. To unequivocally show that the *Kit* allele is not induced in a rare population of cardiomyocytes during development we made *Kit* null embryos by placing the *Kit*-Cre allele over the *Kit*-MerCreMer allele ($Kit^{MCM/Cre}$) with the R-GFP reporter. The rationale for performing this experiment is rather complicated, but it is nonetheless absolutely critical in establishing the fidelity of the Kit allele knock-in strategy for only revealing cells that may have come from the c-kit lineage. Deletion of the *Kit* gene ($Kit^{MCM/Cre}$ mice) renders the heart devoid of functional c-kit^+ progenitor cells, as previously published [23]. So there is little chance to generate $eGFP^+$ cardiomyocytes if they truly came from c-kit^+ cells with progenitor activity, which is exactly what we observed (no $eGFP^+$ cardiomyocytes were present in the embryonic hearts of *Kit* null mice). The "trick" behind the experiment is if we still observed the same low percentage of $eGFP^+$ cardiomyocytes in *Kit* null hearts, it would suggest that they did not come from true c-kit^+ progenitors, but instead likely arose from the Kit allele being turned on in a differentiated cardiomyocyte ectopically (sporadically). Even a little bit of ectopic Cre expression for only a few hours may be enough to permanently label a cardiomyocyte, and without the generation of *Kit* nulls as presented in Figure 4, it would have been very difficult to prove that some or even all of the $eGFP^+$ cardiomyocytes identified in the heart were not the result of some sort of random activation of the *Kit* allele in a rare population of cardiomyocytes for a brief period of time. It is also important to realize that *Kit* null embryos are not disrupted for expression of the Cre or MerCreMer cDNAs from the *Kit* locus, it is simply that the locus no longer makes c-kit protein and hence this renders the progenitor cell population devoid or dysfunctional in the heart, as well as in many other tissues, as shown previously [19-23]. Because deletion of the *Kit* gene resulted in no cardiomyocytes being $eGFP^+$, it essentially proves that the $eGFP^+$ cardiomyocytes observed in the hearts of Kit+/Cre R-GFP mice are due to cellular transdifferentiation (or fusion), and not that the *Kit* allele had become spuriously activated for a brief period of time in a random population of rare cardiomyocytes. Hence, the KI strategy is effective for lineage tracing and the observed cardiomyocytes that are $eGFP^+$ labeled arose from a c-kit^+ cells with some sort of progenitor activity (albeit at very low levels) or due to fusion with a c-kit^+ lineage cell.

Another minor concern is that the levels of new cardiomyocyte formation from the c-kit^+ lineage reported here may be under representative lower due to replacement of one functional *Kit* allele. Thus, there could be less c-kit^+ "progenitor-like" activity in the hearts of these mice due to only a 50% dosage of c-kit protein. However, we did not detect a defect in bone marrow cell numbers associated with heterozygosity of the *Kit* locus or a defect in the status of differentiated hematopoietic cells in adult mice, and mice were otherwise normal suggesting that a 50% reduction in c-kit protein expression did not have a major effect on known sites of progenitor cell activity in the mice (data not shown). Finally, another potential issue is if c-kit^+ cells present in the heart express c-kit at levels below the threshold of the Cre recombinase-based system. If this were the case it would again under-represent the total number of potential cardiomyocytes as being labeled from the c-kit^+ lineage. However, Cre-based lineage tracing is a widely accepted standard for relevant gene expression, so the opposite argument could easily be made that if expression is below the Cre-threshold, it may not be physiologically meaningful. Indeed, our *Kit*-Cre knock-in approach did identify c-kit^+ lineage cells in every cell type and tissue in which it has been previously reported [2,11-13]. Moreover, isolation of total non-myocytes from the hearts of adult $Kit^{+/Cre}$ x R-GFP mice placed in differentiation promoting conditions in culture showed that both $eGFP^+$ and $eGFP^-$ cells could begin to express markers of cardiac differentiation. These data suggest that c-kit-derived cells with one functional *Kit* allele are capable of initiating the

cardiomyogenic program, and that there is not a threshold whereby only lower expressing c-kit cells, which fail to activate the R-GFP reporter system, preferentially give rise to cardiomyocytes.

In summary, the approach employed here definitely shows that endogenous c-kit lineage cells can contribute new cardiomyocytes to the heart, as well as other important cell types (endothelial cells), suggesting that a therapeutic angle could be exploited, as is currently the case in ongoing clinical trials with human patients in heart failure post-MI injury [24]. However, our results also clearly demonstrate that generation of new cardiomyocytes from endogenous c-kit$^+$ cells is a rare event so efforts should be directed at expanding these cells in the heart or in more effectively directing their differentiation towards the cardiomyocyte lineage.

Type of file: table
Label: Supp Table1
Filename: Supp Table1.pdf

Supplementary Table 1. Summary of primary and secondary antibodies used throughout the study

Primary antibodies	Manufacturer	Catalog #		Dilution (paraffin sections)	Dilution (frozen sections)	Dilution (isolated cells)	Dilution (flow cytometry)
sarcomeric alpha-actin	Sigma	A2172	Mouse monoclonal	1:100	1:100	1:300	
alpha-actinin	Sigma	A7811	Mouse monoclonal	1:200	1:100	1:100	
smooth muscle alpha-actin	Sigma	A2547	Mouse monoclonal	1:500			
CD31	eBioscience	17-0311-82	Rat monoclonal				1:100
CD31	BD Bioscience	553370	Rat monoclonal		1:200	1:300	
CD31	LS Bioscience	LS-B4737	Rabbit polyclonal	1:75			
CD3 gamma	LS Bioscience	LS-B7495	Rabbit monoclonal	1:50	1:300	1:300	
CD34	Abcam	Ab81289	Rabbit monoclonal	1:200			
CD34	eBioscience	14-0341-81	Rat monoclonal		1:300	1:300	1:200
CD45	BD Bioscience	550539	Rat monoclonal	1:20		1:300	
CD45	R+D Systems	AF114	Goat polyclonal				1:200
c-Kit	R+D Systems	AF1356	Goat polyclonal		1:100		
c-Kit	BD Bioscience	553356	Rat monoclonal				1:100
c-Kit	BD Bioscience	553353	Rat monoclonal				1:100
Desmin	Millipore	04-585	Rabbit monoclonal	1:100	1:300		
Desmin	Sigma	D1033	Mouse monoclonal	1:100	1:300		
GATA4	Novus Biologicals	NBP1-19461	Rabbit polyclonal			1:200	
GFP	Abcam	Ab290	Rabbit polyclonal	1:200			
GFP	Santa Cruz	sc-9996	Mouse monoclonal	1:50			
NG2	Millipore	Ab5320	Rabbit polyclonal		1:200		
Troponin T	Abcam	ab10214	Mouse monoclonal			1:250	
Vimentin	Abcam	ab45939	Rabbit polyclonal	1:200		1:500	
vWF	Millipore	AB7356	Rabbit polyclonal	1:100			
Mouse Hematopoietic Lineage Biotin Panel	eBioscience	88-7774-75					1:200

Secondary antibodies	Manufacturer	Catalog #		Dilution (paraffin sections)	Dilution (frozen sections)	Dilution (isolated cells)	Dilution (flow cytometry)
Alexa Fluor 405 (streptavidin)	Life Technologies	S32351					1:100

Alexa Fluor 488	Life Technologies	A21141	Goat anti Mouse		1:500	1:500	
Alexa Fluor 488	Life Technologies	A11008	Goat anti Rabbit	1:200	1:500	1:500	
Alexa Fluor 568	Life Technologies	A11011	Goat anti Rabbit	1:200	1:500	1:500	
Alexa Fluor 568	Life Technologies	A11031	Goat anti Mouse	1:200	1:500	1:500	
Alexa Fluor 568	Life Technologies	A11077	Goat anti Rat	1:200	1:500	1:500	
Alexa Fluor 647	Life Technologies	A31573	Donkey anti Rabbit	1:200			
IgG (H+L) - Biotin	Santa Cruz	sc-2042	Donkey anti Goat		1:400		