Peer Review File

Manuscript Title: Morphogen gradient scaling by recycling of intracellular Dpp

Editorial Notes:

Redactions – unpublished data

Parts of this Peer Review File have been redacted as indicated to maintain the confidentiality of unpublished data.

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Romanova-Michaelides et al carefully measure kinetic parameters for the processes associated with the spread of Dpp through the Drosophila wing disc and use these to construct a biophysical model. They interpret the data to conclude that the dominant feature controlling Dpp spread is cell trafficking via internalization by endocytosis and re-exocytosis of ligand molecules. They provide evidence that the rate of endocytosis is increased by Pentagone and this explains how the range of the Dpp scales to the size of the growing wing disc.

The authors are using a well-established model that has become a paradigm for understanding how morphogen gradients are established and scaled in developing tissues. Using this model, different mechanisms of ligand spread and scaling have been proposed by different groups over the last 20 years and debate has continued over the interpretation and validity of different proposals. The current study is an elegant and detailed contribution to the field. By attempting to measure the kinetic parameters associated with the molecular/cellular processes implicated in ligand transport the authors aim to resolve the disputed issues and formulate a consistent model. While I doubt this will end the discussion in the field, the study provides new insight and sets standards for biophysical measurements.
A strength of the study is the combination of biophysical measurements and genetic assays. For example, the evidence that a receptor other than Tkv is responsible for the internalization of Dpp is an important and demonstrates the strength of quantitative assays couple with genetics (Fig 4). Another example, the Dpp-Timer experiments are also very elegant. Moreover, the comprehensive and detailed measurements will be of interest to the wider field.

A difficulty I had with reading the paper is that the short format and presentational style means that it is difficult to assess the rigorousness of the experiments and the confidence one should have in the conclusions. Details of the experiments and key controls are buried in the Methods or Supplemental Figures and often difficult to find. This does not mean I doubt the conclusions, just that it is difficult to assess them from an objective critical perspective. Some examples:
- Does the "quick 4oC acid wash" (line 41) efficiently remove only extracellular GBP-Dendra2?
- How can the authors be confident that the behavior of GBP-Dendra2 bound GFP-DPP is similar to endogenous Dpp?
- Why does there appear to be an inflection in the curve of Nanobody internalization at ~1000s (Fig 2e)? Are there two processes that this is measuring that can explain this. Since this is used to tightly constrain k_n, k_r, k_o: how much would these parameters change if a different value for k_2 is used?
- Was FRAP uniform with the ROI or did it vary in relation to the distance from the edge? How was this accounted for in the data analysis?

I was confused by the interpretation of the photoconversion of internalized Dpp (Fig 1a,b). It appears that the photoconverted region ("above the red line") represents at least some cells producing Dpp. If this is the case, could the intercellular spread of photoconverted Dpp be derived from the secretory pathway, not the endocytic pathway?

The model the authors use to fit their measurements appears to assume no Dpp ligand is lost from the extracellular space in epithelium (e.g. via diffusion out of the basement membrane or into the

peripodial space or via extracellular degradation) - as there is no decay term in eqn S1.1. Is this a realistic assumption? Can loss of Dpp from the wing disc epithelium be quantified and included in the model? How would loss from the epithelium in this way affect the authors' model? Since the apical-basal depth of the epithelium changes during development and the ratio of apical-basal length to anterior-posterior length of the epithelium changes with growth, it is plausible that this could contribute to scaling.
Moreover, the proportion of Dpp in the extracellular space is used to rule out the "the lower cloud of parameter values" in the parameter fit (SM 2.5), indicating the importance of understanding the mechanisms resulting in the extracellular concentration of Dpp.

As well as altering the receptor binding rates, the changes in parameter values the authors calculate for both the smaller discs, pentagone and Dally mutants reduces the Diffusion coefficient substantially. This suggests that pentagone and Dally do more than alter the rate of internalization and recycling. The authors should comment on this and suggest how this can be accommodated in their preferred interpretation.

This group of authors have previous proposed (Wartlick et al 2011 Science) that changes in degradation rate account for gradient scaling in the wing disc. I appreciate the current analysis is more detailed and updates their previous work, but it would be helpful for a reader if the relationship to the previous conclusions were explained.

Referee #2 (Remarks to the Author):

This paper addresses the important question of how morphogen gradients scale on growing domains. The group discovered scaling of the Dpp gradient in the Drosophila wing disc in 2011 (Wartlick et al, Science, 2011). Since then, several scaling mechanisms have been proposed, but key data was missing on the cellular level to test these mechanisms. This paper now uses the latest experimental techniques to provide important data on the cellular basis of scaling. Based on the data, the authors propose that Pentagone mediates gradient scaling on the growing wing disc by changing the diffusion constant, as well as the binding and unbinding rates of Dpp to the internalizing receptor (which they show is not Tkv) because this affects the recycling of the morphogen; the authors provide experimental evidence that part of the internalized Dpp is recycled back to the extracellular space (as had long been assumed in models of gradient dynamics).

It is important to note that most rates can be measured only indirectly, and the inferred values is therefore model-dependent – as directly visible in Extended Data Figure 4, where the parameter values that were inferred with two different models can be compared. Accordingly, the above observations and conclusions are a consequence of the particular modeling assumptions.

The entire analysis in this paper rests on the assumption that the Dpp gradient is in (quasi-)steady state. Even though the initial model that is presented in the Supplementary Material (S1.1-S1.5) is a dynamic model, the authors immediately move to a steady state assumption (which then leads to the characteristic length lambda in Equation 1) without evaluating whether this is warranted. This is surprising given that the data that the authors collected allows such an evaluation for the very first time.

A previous publication (that the authors do not mention) showed that the observed Dpp scaling can directly result from the spreading of the morphogen by diffusion if the gradient does not reach steady state within the patterning phase (Fried & Iber, Nat Commun, 2014). In this case, the gradient lengthens naturally over time, without a need for a change in any of the rate constants. The intuitive explanation for this phenomenon is that diffusion from a source (AP boundary) results in a mean squared displacement of the morphogen that is proportional to time. The diffusion front

can be approximated by an exponential function with a characteristic length lambda that is proportional to the square root of this mean squared displacement, such that the characteristic length lambda is proportional to the square root of time, which is roughly proportional to the square root of the length of the wing disc domain. Quantitative comparison to the scaling data shows that a square root relationship fits the data at least as well as a linear relationship (the red line in Fig. 2a in this paper also seems to follow a square root relationship). In this way, dynamic spreading can explain the observed gradient scaling, and quantitatively reproduces also all the other related experiments.

If the Dpp gradient does not reach steady state during the analysed time window, then the apparent changes in the parameter values that this study infers would only be a consequence of the steady state assumption. In this case, the quasi-steady state assumption would create the conundrum of how rates change over time as the tissue expands, when they are in fact constant.

To judge whether the Dpp gradient is in steady-state, one needs to measure the Dpp half-live. This has indeed been achieved in this study here for the first time (parameter k2), and the reported Dpp half-live is consistent with a pre-steady state expansion of the gradient (Fried & Iber, Nat Commun, 2014). Also, the data on Pentagone would be consistent: a lower koff as observed for the Pent and Dally mutants is sufficient to explain the lack of scaling (Figure 4e, Fried & Iber, Nat Commun, 2014) - without a need for the involvement of further regulatory circuits (expander-repression) that otherwise need to modulate parameter values over time.

The data that the authors generated now allows them to assess whether the pre-steady state nature of the gradient leads to scaling by itself, or whether regulatory feedbacks (expander-repressor) must be invoked. Accordingly, the authors should now use their data to assess whether the gradient is in steady state, or whether their dynamic model (S1.1-S1.5) by itself can quantitatively explain their scaling data. Here, it is important to note that the growth rate is likely to be different from what has been published before (Wartlick et al, Science, 2011) as the Dpp expression levels are now lower. But it should be straightforward to measure this one missing parameter.

Overall, the authors have spent a lot of thought on the analysis, as reflected by the extended Supplementary Material. The experimental data in this study exploits the most recent experimental technologies and presents an important advance over existing data, even if a careful reanalysis may change the conclusion regarding the scaling mechanism.

Suggested improvements in a revision:

Line 182: "This gradient scaling must be mediated by changes in transport rates." This statement is not generally true – scaling is possible also for constant transport rates as long as the gradient is not in steady state (Fried & Iber, Nat Commun, 2014).
Please check the steady-state assumption with the new data, and re-analyse the data with a pre-steady state gradient, if appropriate.
Unlike in the 2011 study, no growth curve appears to be provided. It would be important to provide such a growth rate to be able to solve the model on a growing domain and evaluate scaling; the growth rate is likely lower, given the 10-fold lower Dpp expression in the new system.
The authors state that their estimate of the extracellular fraction hinges on a large decay length limit – again, how large is the error? Does it explain the inferred differences between young and old discs?
In the model description in Supp 1, the domain and with it the values of v_n are not well explained, i.e. for which n is v_n = 1? This is better explained in the preprint, but it would be good to have all information in this Supp text.
Pulse-chase experiments: it is not clear from section 2.4 what was measured (where is the data?) and how it was analysed. I may have missed a more thorough explanation elsewhere in the Supp.

This study provides a lot of data and details. It would be helpful if you could link to it – or add it if it is indeed missing.
FRAP experiments: also, here it is not fully clear how this was analysed.
Supp Section 2.5.2: it is not clear how k_"on" , k_"off" , k_i, D_0 and k_2 were exactly determined – the description is rather vague and no equations are provided.
All analysis scripts and data files should be made available to the reader. Even though the authors provide extensive Supplementary Material, it often would be easier to see what was done if one had access to the code; this would directly address the above unclear points. Also, others can build more easily on the work.
The paper heavily rests on Reference 4, which is listed as published in Phys Rev E in 2019. However, I could only find a preprint https://arxiv.org/abs/1909.13280.

Small Points:

Line 45: Please explain how it follows that photoconverted molecules are only in the endocytic, but not in the secretory pathway. If there is not enough space, please refer to the Supplementary Material.

Figure 1a: add scale bar
Figure 1b: where is zero? Is this the red line in a? How does the amplitude evolve in the part that was photo-converted?
Figure 2f,g: Why was the pent mutant not measured also at a short length?
Figure 2f,g, 3a: it is maybe misleading to present this type of information as bar plot – a box plot would be more appropriate.

The work is spread over five different files (main file, Extended Data, Supp 1, Supp 2, separate manuscript), which makes it hard to follow – but maybe there is no better way.

Referee #3 (Remarks to the Author):

Gonzalez, Jülicher and colleagues combine experimental data and mathematical modeling to propose that scaling of the Dpp morphogen gradient in the growing developing wing disk of Drosophila is achieved by increasing the contribution of the internalized Dpp molecules to the extracellular pool of Dpp. The topic is highly relevant for the developmental biology community and different models have been proposed to address this beautiful question. By combining a good number of newly generated tools with nicely developed and very imaginative experimental settings, authors provide convincing evidence that the pool of Dpp endocytosed is being recycled and sent to the extracellular pool again, that this recycling contributes to gradient formation and that Dpp endocytosis is independent of the Tkv receptor. I have no major issue with this part and as a whole is a very interesting observation. Whether changes in Dpp recycling functionally contributes to scaling and whether pentagone is indeed an important player in morphogen scaling are two relevant questions that are not convincingly addressed in this ms, though. In this context, a recent report of the lab of Arthur Lander in Dev Cell seriously questions the long-range activity of pentagone and rules out a potential role of this gene in Dpp scaling. Indeed, it is proposed in this paper (Zhu et al 2020) that a feedback negative loop between Dpp and its receptors/coreceptors appears to be the one mediating morphogen scaling and it is shown that pentagone activity is rather local and restricted to the lateral regions of the Dpp gradient. The contribution of Dpp recycling to Dpp scaling appears to be rather correlative and its proposed causality is not demonstrated. Authors do provide evidence that Tkv is not mediating Dpp endocytosis. What about Dally and Dlp? Taken all these issues into consideration, I am not convinced that this ms is a serious candidate for Nature.

Referee #4 (Remarks to the Author):

Gradients of diffusible morphogens provide positional information in growing tissues, suggesting that gradients scale with the size of the territory. Deciphering the molecular and cellular basis of scaling is an important issue. Wolpert (1969) first realized that a gradient formed by linear diffusion from a source to a sink scales naturally as the sink moves away from the source. This model, however, requires a local sink and an otherwise stable morphogen. Building upon this idealistic view, Barkai (2010) proposed a model of automatic gradient scaling in which morphogens repress the expression of a diffusible expander that negatively feeds back on morphogen production. In the context of the Dpp gradient in the developing fly wing, a textbook model in Developmental Biology, scaling was proposed to involve the secreted protein Pentagone as an expander of the Dpp gradient. Here, Gonzalez-Gaitan and colleagues study the scaling of the Dpp gradient and propose a related model. In this new model, Pentagone mediates gradient expansion by increasing the rates of Dpp binding to its putative internalization receptor. This is a potentially interesting idea since the secreted protein Pentagone was recently shown to diffuse only minimally (Zhu et al., 2020, Dev Cell, 53, 724-39).

The key findings/claims of this paper (abstract) are that:
- scaling (increase in decay length) is achieved by increased recycling of Dpp
- increased recycling results from changes in the binding rates of Dpp to its internalization receptor
- these binding rates are modulated by Pentagone

Many data (reviewed below) are provided to support these statements. However, to my best understanding, these data provide only indirect evidence in support of the three claims above. Indeed, the rate of Dpp recycling was not measured. Thus, whether this rate changes over time and is modulated by Pentagone remain to be determined. Additionally, the rates of Dpp binding to its receptor were not directly measured. Therefore, whether these binding rates change over time and depend on Pentagone function remain to be established. Of note, the authors propose the existence of a Dpp internalization receptor that is distinct from the signaling receptor Tkv. Because the molecular identity of this putative internalization receptor is not known, analysis of the binding rates cannot be performed.

In this study, the authors first tested whether endocytosed Dpp can contribute to gradient formation. To do so, they used two recycling assays, one based on photoconversion (Fig 1a) and another based on FRAP (Fig 4c,d). Both assays probe the secretion of internalized of GFP-tagged Dpp, i.e. recycling, and measure its diffusion in the target tissue (Fig 1a and 4c,d). The results indicate that Dpp is internalized and recycled and therefore provide support that recycled Dpp contribute to the Dpp gradient (but see points 2, 3 and 6 below). However, the relative contribution of recycled Dpp to the pool of diffusive Dpp that forms the gradient remains to be determined. In other words, these data do not exclude the possibility that the Dpp gradient forms mostly by diffusion of Dpp produced at the source.

Next, the authors proposed a model for the transport of Dpp which involves five pools of Dpp and eight transport steps (Fig 1c). This model considers both diffusive and cell-based transport mechanism. Of note, the model considers Dpp transport by transcytosis but not via cytonemes (i.e. cell size, a, is constant). The model describes the time evolution of Dpp accumulation in the tissue in a set of equations. A decay length of the gradient at steady state can be obtained from these equations. Its value depends on eight independent parameters (transport rate values listed in Extended Table 1). Because most of these transport rate values are not experimentally accessible, the authors developed transport assays that were meant to depend on only a subset of these transport rates (Fig 2; see points 3-6 below for detailed comments on these assays). This approach aimed at constraining the parameter values and at narrowing them down to a limited range of values compatible with the proposed model. In a sense, the experimental data are not meant to challenge the model but to make it work.

Briefly, these transport assays were performed at two time points (l=80 and l=144; see point 4 below) and the results obtained were used to generate a limited sets of compatible parameters values (Fig 3a). Using these parameters, the authors examined the relative contribution of the four modules (Fig 1e) and concluded that transcytosis of Dpp does not contribute to gradient formation (see point 7 below), that diffusion of unbound Dpp is important at early but not late stages, and that recycling of Dpp plays an increasingly important role over time (Fig 3b). Mechanistically, they propose that the binding rates of Dpp to its internalization receptor vary over time.
These conclusions critically depend on the precision and quality of the data from transport assays which were used to estimate Dpp transport rates (see points 2-6). They also depend on assumptions of the model (see point 1 below). Last, since endocytosis and recycling appear to contribute to the Dpp gradient, the authors studied the internalization of GFP-Dpp and confirmed that the Dpp signaling receptor, Tkv, is not required for the internalization of Dpp, suggesting that a Dpp internalization receptor must exist (Fig 4a).

Finally, the authors studied the GFP-Dpp gradient and transport mechanisms in a mutant condition that is known to disrupt scaling. Previous studies showed that the secreted protein Pentagone is required for the scaling of Dpp signaling activity. Here, detailed analysis of pentagone mutant discs indicated that the Dpp gradient does not properly scale in the absence of pentagone (Fig 2a) and that Dpp transport appeared to primarily involved diffusion (Fig 3b). Thus, the authors propose that scaling relies on a time-dependent increase in the concentration of Pentagone, which would lead to a progressive increase in the rate of Dpp recycling. One obvious prediction of this model, and of the original ER model as noted by the authors in p.6 l.229, is that Pentagone levels increase over time and that Pentagone levels correlate with Dpp recycling. This prediction was not examined here. Thus, while it is known that overexpression of Pentagone results in an overexpansion of the Dpp signaling gradient (ref 22), whether this effect is due to changes in Dpp recycling remains to be tested.

In sum, this study addresses a long-standing and debated issue. In order to help the field move forward, robust and precise data are needed if they are to be confronted with the theory. Here, the data primarily serve to provide parameter values for the model. As the main conclusion of this study likely depends on these values, precise quantitative data are critical.

Specific points

1. model
The model looks fine (to the extent I can evaluate it). One assumption of the model, however, may need to be clarified. One main conclusion of this study is that Dpp recycling increases over time, directing scaling. I therefore presume that increased recycling of Dpp at the source results in increased Dpp production. Consistent with this, a previous study by Gonzalez-Gaitan and colleagues found that the amplitude of the DPP gradient increases during development. Despite this, the model appears to consider a constant rate of production from a localized source (SM, Section 1.A.1). If it is indeed so, could the authors justify this choice? And if one considers a variable rate of Dpp production, would the conclusions of the paper remain the same?
Otherwise, the authors should state somewhere that the model best applies to the posterior (P) compartment of the wing disk. Indeed, in the model, cells are either part of the source (vn=1) or not (vn=0), and this parameter remains fixed. This situation corresponds to the P compartment: all P cells are vn=0, and source cells are A cells at the A-P boundary (vn=1). By contrast, the A cells close to the A-P boundary would switch from vn=1 to vn=0 as the tissue grows. This switch is not implemented in the model.
This brings another question. Were all experimental data acquired in P cells? While I presume so, given the legend of Figure 2, it may be better to spell it out clearly. This is important and has several implications. For instance, one might expect that the size of the source of Dpp and of GFP-Dpp driven by Gal4/LexA might slightly differ given the stability of Gal4/LexA which could persist in the A cells that no longer express the dpp gene.

2. Dpp overexpression
A limitation of this study is that Dpp transport rates are estimated using an overexpressed version of Dpp, GFP-Dpp, expressed under the control of dpp-Gal4 or dpp-LexA. Importantly, the authors made the effort to generate an endogenous version of GFP-Dpp using CRISPR/Cas9 knock-in, but this tool had two severe issues (section 2, 1.2.2): first, this dpp allele was homozygous lethal, implying that this version of Dpp is not functional. Second, one copy of GFP-Dpp was too weak to be used in the FRAP assay used to parametrize the model. This reveals that the levels of GFP-Dpp accumulation in the Gal4 (or LexA) conditions are significantly higher than endogenous Dpp levels. Whether this difference matters or not should be examined. Specifically, the authors could study the extent to which endogenous GFP-Dpp and overexpressed GFP-Dpp similarly scale (Ext Data Fig 1a-c), recycle (diffusion rate of internalized and photoconverted Dendra-GBP; Fig 1a,b and Ext Data Fig 1d-f) etc… In other words, it would be important to establish that the values of the Dpp transport parameters do not depend on Dpp levels.
Along the same line, it would be important to state somewhere that most experiments were performed in a context where GFP-Dpp competes with endogenous Dpp. And it is needed to test whether this competition affects the behavior of GFP-Dpp. One possible experiment would be to overexpress untagged Dpp and examine the effect of excess Dpp on the transport rates measured for GFP-Dpp. Since the authors propose that changes in Dpp binding rates underlie scaling, the possibility that competition and overexpression of the GFP reproter affects parameter values should be seriously considered.

3. Dpp recycling and photoconversion
This is a nice assay(Fig 1a,b). It clearly suggests that internalized Dpp can be recycled and diffuse. Since at t=0 post-conversion, a strong signal is seen at the source, recycled Dpp detected at later time points at a distance from the source likely originated from the source cells. However, whether recycling also occurs in the growing target tissue, which is critical for the spreading of Dpp, is assumed but not demonstrated here.
More importantly, it is not clear whether recycling is actually important for the formation of the Dpp gradient. Is it possible to experimentally up(down)-regulate the recycling of Dpp and examine the consequences on the Dpp gradient? (I must say that I have no good suggestion to offer here)
Otherwise, I do not understand how averaging intensity values in three independent experiments/discs (n=3) can produced high amplitude variations along the x-axis (see t=87.5min data plots). Please comment.

4. scaling and data selection
The data shown in Fig 2a are both nice and important. In this plot, the decay length values are given for different binning of length values. This is unusual. Would it be possible to plot instead the raw data as a cloud of individual data points? I am asking this because I did not understand why the authors chose to bin the decay values in two categories corresponding to 144.6 +/-4 and 79.3 +/-1.4 (see legend of Fig 2a; note that different SEM values are given in Ext Data Fig 1b and Table 2; please clarify). The motivation for this very specific binning remains unexplained. Specifically, I am concerned that the corresponding decay length values (28.1 +/-1 and 1.9 +/-0.5; Table S2) may not fit well on a linear regression of the data plotted in Fig 1a, suggestive of data selection. For this reason, I would strongly recommend avoiding categorization whenever possible. I do understand, however, that categorization cannot be avoided when performing transport assays. However, the size distribution of the experimental discs needs to be reported. Currently, this information is not easy to find (or not provided). For instance, in Fig 2e, the authors used l=144 simply to name a sample category. This naming cannot replace the mean/sem values corresponding to this particular sample. Please provide this information (in this and all other experiments).

5. nanobody internalization assay
This nice assay measures the rate of endocytosis of GFP-Dpp bound to a nanobody, following a binding step at 4°C (note: please state in section 2, 2.2.2, that discs are incubated with the same concentration of GBP-Alexa555 during imaging at 25°C, hence allowing for fluid phase

endocytosis; if GBP-Alexa555 is not added, fluid-phase endocytosis may be underestimated).
While this assay aims at quantifying endocytosis at steady-state, it actually measures a pulse of Dpp internalization following transfer of the disc at 25°C. This is an issue. Indeed, after 30 min of incubation, rapid Dpp internalization is no longer detected and only fluid-phase endocytosis is monitored for another 30 min (the linear regime of internalization corresponds to fluid-phase endocytosis, as shown in Ext Data Fig 7i). In sum, this assay may not monitor Dpp endocytosis at steady state.
Surprisingly, a 5-fold variation in fluid-phase endocytosis (phenomenological parameter A in Ext Data Fig 1f) was observed between early and late discs (Ext data Table 2: 0.0011 vs 0.00021). This 5-fold variation likely reflects a combination of technical and biological noises.
On a related note, the accumulation profile of internalized nanobody appeared somewhat variable (compare Figs 2c, 4b and 4d): is variability due to biological or technical variations?
Otherwise, what are the samples (n=15) shown in Ext Data Fig 1d?

6. iFRAP
FRAP was performed on the internalized pool of GBP-Alexa555 (Fig 4c-g). Internalization, following a 4°C binding step, was performed for 35 min, ensuring that a large fraction of internalized GBP-Alexa555 entered in GFP-Dpp dependent manner. Following photobleaching, fluorescence recovery was monitored over the next 30 min. Importantly, at this stage (35-60min of incubation), GBP-Alexa555 was reported to be mostly endocytosed in a fluid phase manner. Thus, whether this assay monitors the receptor-dependent endocytosis of a GFP-Dpp/GBP-Alexa555 complex remains to be established. It is therefore unclear whether this assay actually monitors the recycling, diffusion and endocytosis of GFP-Dpp.

7. Transcytosis
Previous work by Gonzalez-Gaitan has suggested that transcytosis is required for the formation of the Dpp gradient (Entchev et al., Cell, 2000). Here, Gonzalez-Gaitan suggests that transcytosis plays no significant role (Fig 3b). Given that the original finding was very controversial, this new conclusion by the same authors should more clearly appear in the main text. Currently, a discussion on the possible role of transcytosis is buried in section 2, 4.1.3, and will only be accessible to a very limited audience.

8. DppTimer
The green/red ratio is interpreted to reflect the age of Dpp. However, this ratio also depends on intracellular pH. Following internalization into acidic endosomal compartments, the fluorescence signal of sfGFP (but not those of mKate2) is quenched. Thus, this approach cannot be used to directly infer the ages of the Dpp molecules.

9. Pentagone function
The data shown in Ext Data Fig 1c suggests that the Dpp gradient cannot be approximated to a simple exponential. Is this observation reproducible? Could the authors quantitatively examine how well the experimental data fit with an exponential in both wild-type and pentagone mutant discs? In other words, does the loss of Pentagone affect the shape of the Dpp gradient? This needs to be examined.

Author Rebuttals to Initial Comments:

Referee #1 (Remarks to the Author):

Romanova-Michaelides et al carefully measure kinetic parameters for the processes associated with the spread of Dpp through the Drosophila wing disc and use these to construct a biophysical model. They interpret the data to conclude that the dominant feature controlling Dpp spread is cell trafficking via internalization by endocytosis and re-exocytosis of ligand molecules. They provide evidence that the rate of endocytosis is increased by Pentagone and this explains how the range of the Dpp scales to the size of the growing wing disc.

The authors are using a well-established model that has become a paradigm for understanding how morphogen gradients are established and scaled in developing tissues. Using this model, different mechanisms of ligand spread and scaling have been proposed by different groups over the last 20 years and debate has continued over the interpretation and validity of different proposals. The current study is an elegant and detailed contribution to the field. By attempting to measure the kinetic parameters associated with the molecular/cellular processes implicated in ligand transport the authors aim to resolve the disputed issues and formulate a consistent model. While I doubt this will end the discussion in the field, the study provides new insight and sets standards for biophysical measurements.

A strength of the study is the combination of biophysical measurements and genetic assays. For example, the evidence that a receptor other than Tkv is responsible for the internalization of Dpp is an important and demonstrates the strength of quantitative assays couple with genetics (Fig 4). Another example, the Dpp-Timer experiments are also very elegant. Moreover, the comprehensive and detailed measurements will be of interest to the wider field.

A difficulty I had with reading the paper is that the short format and presentational style means that it is difficult to assess the rigorousness of the experiments and the confidence one should have in the conclusions. Details of the experiments and key controls are buried in the Methods or Supplemental Figures and often difficult to find. This does not mean I doubt the conclusions, just that it is difficult to assess them from an objective critical perspective. Some examples:

- Does the “quick 4oC acid wash” (line 41) efficiently remove only extracellular GBP-Dendra2?

We already addressed this issue for the GBP-Alexa555 nanobody, but indeed not for GBP-Dendra2 which only differ in the fluorophore (Extended Data Figure 5b,c, 8a-d). Prompted by this question of the reviewer, in the revised version we show

a) that the acid wash does *“efficiently remove”* extracellular GBP-Dendra2, so that we can exclude the possibility that the recovery of the gradient in the target area occurs at the expense of extracellular GBP-Dendra2 molecules at low, undetectable concentrations (below the detection limit), which could have in principle remained after the acid wash.

b) to what extent the acid wash “*remove only extracellular GBP-Dendra2*” *versus* intracellular molecules.

a) To address the efficiency of the acid wash for GBP-Dendra2, we incubated the disc with GBP-Dendra2 at 4˚C, which labels extracellular GFP-Dpp, followed by acid wash, which would in principle remove all the extracellular label. After raising the temperature, we then checked whether internalization in endosomal structures is observed at the expense of a potential extracellular low level pool remaining after acid wash. Such internalization is not observed (NEW Extended Data Figure 5e,f), indicating that acid wash of extracellular GBP-Dendra2 is efficient and the recovery of the gradient in the photoconversion experiment (Fig. 1a-c) is due to the mobilization of intracellular molecules in the photoactivated area. This type of control experiment was already shown in the previous version of the manuscript for GBP-Alexa555 and FRAP and is confirmed now in the revised version for GBP-Dendra2 and photoconversion.

b) In the photoconversion experiment, it is essential that extracellular molecules are removed such that we can be sure that the recovery is not due to residual extracellular molecules. This is now controlled as described in point a) above. In contrast, it is not very important that intracellular pools are also affected to some extent. Even so, we confirmed that the intracellular pool is not substantially affected (NEW Extended Data Figure 8e). Upon acid wash, fluorescence of the intracellular pool decreases only by $2.3 \pm 0.6\%$. This was expected, because the incubation (and the wash) is performed at 4˚C (no endocytosis can occur) and without any permeabilization, and therefore the acid wash can only remove the extracellular space and not intracellular compartments.

We now added these new controls for the acid wash in the Supplementary Material (section 2.7,2.8) and in Extended Data Figures 5 and 8.

- How can the authors be confident that the behavior of GBP-Dendra2 bound GFP-DPP is similar to endogenous Dpp?

To study the behavior of GFP-Dpp bound to GBP-Dendra2, we studied the steady state decay length of GFP-Dpp^{LOP} without the nanobody in comparison to the recovered gradient of GBP-Dendra2* bound to GFP-Dpp in the photoconversion experiment. The decay lengths are not different, indicating that the trafficking properties of these molecules are the same to the extent that they can generate the same gradient properties (NEW Extended Data Figure 5g).

- Why does there appear to be an inflection in the curve of Nanobody internalization at ~1000s (Fig 2e)? Are there two processes that this is measuring that can explain this.

In the previous version of the manuscript, in the methods chapter on the nanobody experiment (Section 2.2.1 of Supplementary Material; see also Extended data Fig. 1f,g), we show how these two regimes in the dynamics of nanobody internalization (separated by the "inflection point" mentioned by the reviewer) emerge from the internalization, recycling and degradation dynamics. More precisely, it can be shown that the nanobody internalization dynamics $C_T(t)$ is captured by the following phenomenological equation (Supplementary Material Section 2 chapter 2.2.1, Extended data Fig. 1f,g),

$$C_T(t) = \mathrm{C}(1 + B(1 - e^{-pt}) + At)$$

Where C is proportional to the nanobody concentration and B, p and A are phenomenological parameters, that describe the dynamics of the nanobody internalization. This equation reveals two regimes: 1) an exponential part ($1 + B(1 - e^{-pt})$, red box in Extended data fig. 1f,g) defined by the prefactor B and the relaxation time p and 2) a linear part (At, green box in Extended data fig. 1f,g) defined by the slope A.

These 2 regimes (exponential and linear) are expected in the nanobody internalization dynamics. Indeed, solving the general dynamic equations of the system (equations S1.1-S1.5) for the particular case of the nanobody experiment (equation S2.14 from supplementary material), we show that

$$C_T(t) = C(1 + \frac{k_N k_r}{(k_o + k_r)^2}\left(1 - e^{-(k_r + k_o)t}\right) + \frac{k_N k_o}{k_o + k_r}t)$$

These two equations have the same structure: an exponential plus a linear regime. Therefore, the parameters B, p and A depend on the internalization rate k_N, recycling rate k_r and the output rate from the mobile pool k_o as follows:

$$B = \frac{k_N k_r}{(k_o + k_r)^2}$$

$$p = (k_r + k_o)$$

$$A = \frac{k_N k_o}{k_o + k_r}$$

The relaxation dynamics of the exponential regime is indeed captured by the rate $p = (k_r + k_o)$, which depends on the recycling rate k_r and the rate k_o of output from the mobile endosomal pool. Related to the saturation level of this exponential regime, the parameter B also depends on k_r and k_o and, in addition, it depends on the endocytosis rate k_N (see derivation in section 2.2.1). The exponential regime does not relax to a constant saturation level in steady state, but it relaxes into a linear regime. This is because the fluorophore of the nanobody (the synthetic molecule Alexa555) cannot be degraded in the lysosome in the time scale of the experiment (see Extended Data Figure 7i-l): as a consequence, the fluorophore accumulates linearly in the assay. Therefore, the nanobody internalization assay does not capture either of the two degradation rates k_1 and k_2. The linear regime is characterized by the slope A that depends on k_r, k_o and k_N (see above).

Indeed, the observation that the internalization assay reveals two regimes (Fig. 2e) is predicted in our theory and validates the phenomenological transport description that underlies this work. A detailed derivation can be found in the Supplementary Material Section 2 chapter 2.2.1. We have now added a clarification about how the two dynamics spotted by the reviewer emerge from the different cellular rates in the legend of Fig. 2e.

Since this is used to tightly constrain k_n, k_r, k_o: how much would these parameters change if a different value for k_2 is used?

As explained above, the dynamics of this assay do not depend on k_2. Therefore, the values of k_r, k_o and k_N would not change for different values of k_2. For details, see 2.2.1.

- Was FRAP uniform with the ROI or did it vary in relation to the distance from the edge? How was this accounted for in the data analysis?

FRAP recovery is not affected by the distance of the ROI to the edge of the source. Indeed, in a previous report[1], we already showed that D_{eff} and k_{eff} obtained from FRAP assays are robust and do not change for ROIs at different distances from the source boundary (the edge of the source domain), see Supplementary figure S4 from Kicheva *et al.*, 2007[1].

Within the ROI, during recovery, a small proximo-distal gradient is expected as revealed by a new theoretical study which we include in the revised version (see NEW Extended Data Figure 8m). The spatial differences within the ROI are however too small to be detected considering the inherent noise in the FRAP data. This is fully accounted for in the estimation of parameters: the theoretical dynamics are computed from the average within the simulated bleached region and the experimental recovery dynamics is determined as the average in the bleached ROI. In the supplementary material describing the FRAP assay (section 2.5), we included now this consideration.

I was confused by the interpretation of the photoconversion of internalized Dpp (Fig 1a,b). It appears that the photoconverted region ("above the red line") represents at least some cells producing Dpp. If this is the case, could the intercellular spread of photoconverted Dpp be derived from the secretory pathway, not the endocytic pathway?

Because of the way the experiment is done, it can be excluded that the origin of the photoconverted molecules that spread is exclusively secretory. This is because, in the experiment, GBP-Dendra2 nanobodies are first internalized and the extracellular pool of nanobodies is eliminated by an acid wash which is efficient (see above). At this point only endocytosed GBP-Dendra2 molecules remain in the tissue and it is only then that Dendra2 is photoconverted: This internalized pool is the only source of molecules that can spread in the tissue. Therefore the origin of GBP-Dendra2 cannot be exclusively secretory.

It is however possible in principle that a part of the internalized pool traffics through a secretory compartment. In this scenario, the route for Dpp recycling would involve the secretory pathway (for which there are some examples in the literature[2,3]), which is indeed a potential mechanism of recycling both in the source domain and in the target. The alternative to this route involving the secretory pathway involves recycling endosomes[4].

In the revised version, we eliminated the strict exclusion of the secretory pathway as a possibility:

> *page 2. Because the photoconverted molecules were only in the endocytic pathway, ~~and not in the secretory pathway~~ this result indicates…*

This scenario of endocytic/secretory mixage would only be an issue if the two pathways can mix only in the source domain, but cannot mix in the target tissue. In this hypothetical scenario, the photoconversion of internalized GBP-Dendra2 in the source domain does not exclude that, in the target cells, endocytosed molecules may not be re-externalized again. To eliminate this possibility, and show that endocytosed molecules are re-externalized from target cells, we performed a new series of experiments where we incubated with GBP-Dendra2 and performed acid wash as before, followed now by photoconversion in a ROI in the target tissue, which does not overlap with the source domain. The NEW Figure 1b shows that endocytosed GBP-Dendra-Dpp molecules do spread from target cells and can form a gradient. In the main text of the revised version of the manuscript we added:

> *page 2. Photoconversion of this intracellular GFP-Dpp-bound GBP-Dendra2 in a region of interest (ROI; whether in the source (Fig. 1a) or target tissue (Fig. 1b)) allows us to address whether these internalized molecules can move to neighbor cells and form a gradient there (for details and controls of this assays, see SM 2.8; Extended Data Figures 5e-g, 8a-f).*

and

> *page 2. Because photoconversion in the target territory also lead to formation of a gradient (Fig. 1b), this shows that Dpp molecules can be re-exocytosed also from the target cells.*

The model the authors use to fit their measurements appears to assume no Dpp ligand is lost from the extracellular space in epithelium (e.g. via diffusion out of the basement membrane or into the peripodial space or via extracellular degradation) - as there is no decay term in eqn S1.1. Is this a realistic assumption?

Can loss of Dpp from the wing disc epithelium be quantified and included in the model? How would loss from the epithelium in this way affect the authors' model? Since the apical-basal depth of the epithelium changes during development and the ratio of apical-basal length to anterior-posterior length of the epithelium changes with growth, it is plausible that this could contribute to scaling.

Moreover, the proportion of Dpp in the extracellular space is used to rule out the "the lower cloud of parameter values" in the parameter fit (SM 2.5), indicating the importance of understanding the mechanisms resulting in the extracellular concentration of Dpp.

The reviewers raise two related issues: leakage from the extracellular space and extracellular degradation.

Leakage from the extracellular space. As noted by the reviewer, we have not ourselves considered the possible effects of leakage of molecules out of the epithelium. However, a recent study which explores the possible effects of molecular leakage in a synthetic system in the fly wing provides interesting insights[5]. From the results of this study we conclude that the formation of the Dpp gradient *in vivo* is unlikely to be influenced by loss of molecules in the extracellular space.

Indeed, both an apical and a basolateral gradient were quantified in that report[5] (their Figure S3D). That article concludes that leakage is only relevant for the basolateral gradient: i) the basolateral gradient shows an offset which reflects the leakage-in of molecules, but the apical gradient does not show such offset (their Figure S3D) and ii) the fat body trap does not affect the apical gradient (their Figure S3C compared to Figure 1D). Therefore no detectable leakage of molecules was found in the apical gradient.

Indeed, upon expression of a secreted form of GFP in the Dpp source region, they observed that it can form a gradient in the target that differs in the apical and basal region: in the basal region there is a prominent offset (as large as the gradient amplitude) which reflects the fact that molecules leaked from the tissue into the hemolymph can return back to the tissue in the form of an offset. There is no offset in the apical gradient, indicating that there is no leakage from/to the apical gradient. In addition, there are tenfold more GFP molecules in the apical than in the basal gradient: the basal gradient is only a very minor component. In this report, we consider only the GFP-Dpp gradient in the apical side (the apical-most 5 µm in the epithelium), since indeed we have previously shown that the basal component is minor (see Kicheva *et al.*, 2007[1] and unpublished). This is now considered in our discussion of previous reports in the literature (Supplementary Material Section 2 chapter 4.3).

As suggested by the reviewer, we attempted to estimate the loss of extracellular molecules by leakage: we labelled extracellular GFP-Dpp with GBP-Alexa555 and chased the living disc for 7 hours at 4˚C (where no production, internalization or degradation happens) to determine how fast could Dpp molecules leak from the extracellular space (NEW Extended Data Figure 13d,e). We show that the decay is negligible, indeed consistent with a scenario in which leakage is not important in the Dpp system.

Based on these arguments, a significant leakage rate of Dpp *in vivo* is unlikely to take place which justifies our assumption of negligible ligand loss in the extracellular space. As requested by the reviewer, we still studied the theory including our five Dpp pools and cellular rates and allowing the unbound pool to be affected by leakage (Supplementary Material Section 2 chapter 3.10). We then studied what would be the consequences of leakage on the steady-state gradient.

To clarify this issue, we wrote in the main text of the revised manuscript:

> *page 2. Unbound ligand molecules could in principle leak out of the epithelium, as recently studied in a synthetic system where secreted GFP binds to a membrane nanobody and forms gradients[5]. The gradients are different in the apical and basal side of the epithelium. Basally, the gradient is fainter than apically and shows a large offset (as large as the amplitude of the gradient). This seems to be due to GFP molecules which leaked out to the hemolymph and back again to the epithelium. In the same discs, the apical gradient has a ten-fold larger amplitude and does not show such offset, indicating that leakage from the apical side is not important for formation of this GFP gradient. Because most GFP-Dpp molecules are present in the apical-most 5 µm of the epithelium[1], where there is no leakage, we did not consider leakage as a major phenomenon in the system (see further considerations in SM 2.11, SM 3.10).*

Extracellular degradation: Some extracellular proteins can be cleaved outside cells. In fact, there is no indication of extracellular cleavage of Dpp itself (or any other BMP ligand in Drosophila or other species). Dpp is matured by Furin cleavage of a precursor in the trans-golgi network and not the extracellular space. To date no proteases are known that can inactivate Dpp (or other BMPs) by cleavage extracellularly. Dpp signaling in the embryo is controlled by cleavage of the Dpp-binding protein Sog (but not Dpp itself) by the Tld protease. In any case, during imaginal disc patterning and growth in the larva, this machinery does not function in Dpp spreading and signaling: mutants in this extracellular pathway (sog, Tsg, Tld, etc), which operates during dorso-ventral patterning in the embryo, do not show a Dpp signaling phenotype during larval phase and show only a later pupal function during the differentiation of veins. Therefore, there is no indication that extracellular cleavage of Dpp takes place in the wing.

As well as altering the receptor binding rates, the changes in parameter values the authors calculate for both the smaller discs, pentagone and Dally mutants reduces the Diffusion coefficient substantially. This suggests that pentagone and Dally do more than alter the rate of internalization and recycling. The authors should comment on this and suggest how this can be accommodated in their preferred interpretation.

To be clear, we showed that in the three conditions (small discs, *pent* and *dally* mutants), the contribution of the recycling **MODULE** to the decay length is smaller than in large discs at the expense of the unbound **MODULE** (see Fig. 3b and Extended Data Figure 6f), and this happens without an effect on the internalization or recycling **RATES**, but by changing the extracellular diffusion coefficient (D_0) and the binding/unbinding rates ($\mathrm{k_{on}}$, $\mathrm{k_{off}}$) to the receptor which internalizes Dpp (see Fig. 3a). There is a chapter (3.1) in the Supplementary Material section 2 explaining this difference between the recycling module and the recycling rate. Moreover, to clarify this point, in the new version of the main text of the manuscript we wrote:

> *page 3. Note that the recycling module is not equivalent to the recycling rate, as it depends on all the transport rates, except k_2. Importantly, extracellular diffusivity and the binding rate are essential for the three modules described above.*

Therefore, the actual internalization and recycling rates are not affected in these conditions. Only D_0, $\mathrm{k_{on}}$ and $\mathrm{k_{off}}$ rates are considerably changed. In other words, the contribution of the recycling module does not change due to changes in the recycling rate (or any other intracellular rate, including the internalization rate). Instead, it is modulated *via* changes in the extracellular rates (D_0, $\mathrm{k_{on}}$ and $\mathrm{k_{off}}$) that make intracellular molecules available to return to the extracellular unbound pool and diffuse to spread and contribute to a longer decay length of the gradient.

The reviewer asks to comment on what causes D_0, $\mathrm{k_{on}}$ and $\mathrm{k_{off}}$ to change. This was briefly explained in the previous version of the manuscript and we make this clearer in our revised version:

> *page 8. Dally binds Dpp[6] and it can act on Dpp by facilitating its diffusion in the extracellular matrix milieu[7]. This could explain how Pentagone/Dally influence D_0. Furthermore, Dally is internalized[8-13] and can therefore act as one of the internalization receptors which mediates Dpp endocytosis. We speculate that, if Dally does function as one of the receptors which internalize Dpp, then the $\mathrm{k_{on}}$ and $\mathrm{k_{off}}$ rates that are regulated by Pentagone include the binding/unbinding rates of Dpp to Dally (for further discussion on how Pentagone/Dally could modulate D_0, k_{on} and k_{off}, see SM 4.5.2-4.5.4).*
>
> *Our work suggests that Pentagone mediates scaling by modifying the binding/unbinding properties of Dpp, both for its internalization receptor (possibly Dally) and the extracellular matrix. Pentagone modulates these extracellular parameters which in turn control the importance of recycling for transport. During scaling, Pentagone serves as a clutch to engage the recycling module as the tissue grows to drive the system between transport regimes.*

In the previous version of the manuscript, in Supplementary Material Section 2 chapters **4.5.2** (Dpp binding and unbinding rates ($\mathrm{k_{on}}$ and $\mathrm{k_{off}}$) and extracellular diffusivity (D_0), **4.5.3** (How is D_0 changed by Dally and Pentagone) and **4.5.4** (How does Pentagone and Dally affect the values of $\mathrm{k_{on}}$ and $\mathrm{k_{off}}$?), we speculated on the mechanisms of how Dally and Pentagone modulate the three extracellular rates. In the new version, for clarity, we refer explicitly to these explanations from the main text.

Briefly, Pentagone/Dally can affect the extracellular diffusivity of Dpp by hampering the movement of Dpp in the extracellular space through binding to molecules in the extracellular space (either the extracellular matrix or binding partners in the plasma membrane which are not involved in Dpp internalization). Indeed, we showed in 4.3.2 how the effective extracellular diffusivity (D_0; the parameter we measure through our assays) depends on the true diffusivity of unbound, free Dpp molecules in the extracellular space ($\mathrm{D_f}$) but also on the binding/unbinding of Dpp to extracellular molecules ($\mathrm{K_{on}}$ and $\mathrm{K_{off}}$, with capital K to distinguish it from the binding to the internalization receptor(s)). This is captured by the expression $D_0 = \frac{D_f}{1+K_{on}/K_{off}}$ (equation S2.60). We speculate that Dally could be the internalization receptor and Pentagone changes the $\mathrm{k_{on}}$ and $\mathrm{k_{off}}$ of binding of Dpp to Dally. Therefore, we speculate that the three extracellular rates are modulated by changing the binding/unbinding of Dpp to the internalization receptor and binding partners in the extracellular space.

This group of authors have previous proposed (Wartlick et al 2011 Science) that changes in degradation rate account for gradient scaling in the wing disc. I appreciate the current analysis is more detailed and updates their previous work, but it would be helpful for a reader if the relationship to the previous conclusions were explained.

As the reviewer points out, in the Wartlick *et al.*[14] paper, we showed that scaling of the gradient is based on changes of the EFFECTIVE degradation. We confirmed this observation in this new report by measuring the effective degradation in FRAP experiments (see Extended Data Fig. 8l). This confirmation and the comparison of our new data with the data in Wartlick *et al.* is explicitly considered in chapter 3.3 of the Supplementary Material (Section 2). The two sets of FRAP data align.

In chapter 3.3 ("Effective rates of Dpp transport") we explain what is this effective degradation and how it relates to the trafficking rates:

> *The morphogen gradient range (i.e. its decay length), depends on the effective diffusivity of the morphogen and its effective degradation rate as the molecules spread in the tissue[1,15,16].*
>
> *These effective rates characterize the apparent diffusion and degradation of the total pool of morphogen molecules considered at the tissue length scale. The apparent diffusion and degradation depend only in part on the Brownian motion of molecules or their lysosomal degradation at the cellular length scale, and are also affected by other transport phenomena[1,15,16].*
>
> *The effective diffusivity and effective degradation at the tissue level has been previously estimated by fitting the dynamics of FRAP recovery using a single dynamic equation[1]. With a single equation, one does not distinguish between the dynamics of the different pools, but instead describes the effective behavior of the total pool of Dpp molecules at the tissue level. Using this procedure, our FRAP data yields an effective diffusivity* $D_{eff} \simeq 0.36 \pm 0.02\ \mu m^2 s^{-1}$ *for large discs (Extended Data Fig. 8l, Table 2). Interestingly, this effective diffusivity* D_{eff} *is three-fold smaller than the diffusivity* D_0 *of unbound Dpp molecules estimated above (Fig. 3a). The effective diffusivity and extracellular diffusivity are conceptually different, and this is confirmed here by our measurements.*

Since the reviewer makes us realize that this may lead to some confusion, for clarity, we explicitly call within the main revised text all these considerations which appeared in the Supplementary Material:

> *page 4. iii) Fluorescence recovery after photobleaching in target tissues (FRAP) in a new primary culture system (Extended Data Fig. 1i,j; SM 2.5), which confirms and extends previous FRAP analysis[1,14]; for detail comparison see SM 3.3-3.4; Extended Data Fig. 8l);*

Referee #2 (Remarks to the Author):

This paper addresses the important question of how morphogen gradients scale on growing domains. The group discovered scaling of the Dpp gradient in the Drosophila wing disc in 2011 (Wartlick et al, Science, 2011). Since then, several scaling mechanisms have been proposed, but key data was missing on the cellular level to test these mechanisms. This paper now uses the latest experimental techniques to provide important data on the cellular basis of scaling. Based on the data, the authors propose that Pentagone mediates gradient scaling on the growing wing disc by changing the diffusion constant, as well as the binding and unbinding rates of Dpp to the internalizing receptor (which they show is not Tkv) because this affects the recycling of the morphogen; the authors provide experimental evidence that part of the internalized Dpp is recycled back to the extracellular space (as had long been assumed in models of gradient dynamics).

It is important to note that most rates can be measured only indirectly, and the inferred values is therefore model-dependent – as directly visible in Extended Data Figure 4, where the parameter values that were inferred with two different models can be compared. Accordingly, the above observations and conclusions are a consequence of the particular modeling assumptions.

There might be a misunderstanding here about Extended Data Figure 4. In this figure we are not parameterizing using experimental data. Instead, in this figure, we are considering theoretically how, within the same phenomenological framework, both a model of pure extracellular diffusion and of transcytosis (the two prevailing models in the field) can be captured depending on the choice of parameter values. Those parameters are compatible with the state of the art experimental data before this report, motivating the need to develop new assays, as established in this report, to clarify which model (the two ones proposed until now or a new one) describes Dpp transport *in vivo*. Therefore in Extended Data Figure 4 we do not infer parameters, but show in a theoretical study, that different models are consistent with a single phenomenological framework. Our observations and conclusions are not the consequence of particular modeling assumptions.

Indeed, our approach is not dependent on a particular model, but instead consists of developing a phenomenological description that takes into account processes that we know occur in the tissue: extracellular diffusion, binding, unbinding, endocytosis, degradation, etc. Within this approach, different transport models correspond to certain parameter regimes within the same common phenomenological description. In particular, the two models of free extracellular diffusion and of pure transcytosis correspond to extreme and opposite cases

which can both be captured within our phenomenological framework. Therefore, the inferred values of the parameters of the phenomenological description are not model dependent. We do not “infer parameters with two different models” as the referee suggests. Therefore, our conclusions are not a consequence of particular modeling assumptions but are uniquely obtained in a single phenomenological framework. This description can capture a wide variety of models (not just one or two favoured ones) corresponding to different parameter regimes which reflect transport scenarios ranging from an extreme extracellular diffusion model (observed in *pent* and *dally* mutants) to a recycling model (in large discs) or an intermediate model (in small discs, where both the unbound module and the recycling module are equally important; see Extended Data Figure 4).

In the new version, we now emphasize that we do not infer parameters using two different models, which would depend on model assumptions. Instead we use an unbiased phenomenological description that takes into account trafficking processes by sets of trafficking modules. We compare our results to behaviors found in different models, which are special limiting cases of our general approach. In two locations of the main text of the revised version we wrote:

> *page 2. This description is equanimous: depending on parameter values, the same general description allows to capture a wide variety of transport models. For example, within this description, the recycling rate or the unbinding rate might be negligible, leading to a scenario where only extracellular molecules may contribute to the formation of the gradient (for examples of extreme scenarios, see Extended Data Figure 4).*

> *page 5. To characterize the Dpp transport regime in large discs, we estimate the contributions of four transport modules from equation 1 using the measured rate values. A particular regime is characterized by the prevalence of one or several transport modules over others. Examples of regimes include scenarios where either the unbound module, the transcytosis module or the recycling module prevail (Supplementary Material section 4.1.2). Scenarios with mixed modules are also possible. The relative contribution of a module depends in turn on the actual values of the transport rates (see above). It is worth noting that this approach is equanimous: all possible regimes can be described by the same description of transport (Fig. 1d) and the actual regime emerges from the set of rate values determined experimentally in our approach (Fig. 3a). Therefore, a regime is not a consequence of particular modeling assumptions, but emerges from parameter values determined experimentally.*

The entire analysis in this paper rests on the assumption that the Dpp gradient is in (quasi-)steady state. Even though the initial model that is presented in the Supplementary Material (S1.1-S1.5) is a dynamic model, the authors immediately move to a steady state assumption (which then leads to the characteristic length lambda in Equation 1) without evaluating whether this is warranted. This is surprising given that the data that the authors collected allows such an evaluation for the very first time.

A previous publication (that the authors do not mention)

We apologise for this oversight. We did refer to this report in a previous version of the manuscript, but it must have disappeared by mistake in a late edition of the text. We overlooked this before submission. The reference is added in the new version of the manuscript.

showed that the observed Dpp scaling can directly result from the spreading of the morphogen by diffusion if the gradient does not reach steady state within the patterning phase (Fried & Iber, Nat Commun, 2014). In this case, the gradient lengthens naturally over time, without a need for a change in any of the rate constants. The intuitive explanation for this phenomenon is that diffusion from a source (AP boundary) results in a mean squared displacement of the morphogen that is proportional to time. The diffusion front can be approximated by an exponential function with a characteristic length lambda that is proportional to the square root of this mean squared displacement, such that the characteristic length lambda is proportional to the square root of time, which is roughly proportional to the square root of the length of the wing disc domain. Quantitative comparison to the scaling data shows that a square root relationship fits the data at least as well as a linear relationship (the red line in Fig. 2a in this paper also seems to follow a square root relationship). In this way, dynamic spreading can explain the observed gradient scaling, and quantitatively reproduces also all the other related experiments.

If the Dpp gradient does not reach steady state during the analysed time window, then the apparent changes in the parameter values that this study infers would only be a consequence of the steady state assumption. In this case, the quasi-steady state assumption would create the conundrum of how rates change over time as the tissue expands, when they are in fact constant.

To judge whether the Dpp gradient is in steady-state, one needs to measure the Dpp half-live. This has indeed been achieved in this study here for the first time (parameter k2), and the reported Dpp half-live is consistent with a pre-steady state expansion of the gradient (Fried & Iber, Nat Commun, 2014). Also, the data on Pentagone would be consistent: a lower koff as observed for the Pent and Dally mutants is sufficient to explain the lack of scaling (Figure 4e, Fried & Iber, Nat Commun, 2014) - without a need for the involvement of further regulatory circuits (expander-repression) that otherwise need to modulate parameter values over time.

The data that the authors generated now allows them to assess whether the pre-steady state nature of the gradient leads to scaling by itself, or whether regulatory feedbacks (expander-repressor) must be invoked. Accordingly, the authors should now use their data to assess whether the gradient is in steady state, or whether their dynamic model (S1.1-S1.5) by itself can quantitatively explain their scaling data.

Here, it is important to note that the growth rate is likely to be different from what has been published before (Wartlick et al, Science, 2011) as the Dpp expression levels are now lower. But it should be straightforward to measure this one missing parameter.

The potential non-stationarity of the system is an interesting and important point raised by this reviewer; we took it very seriously and performed a number of new experiments and theory: 1) we measured again the growth rates during development, 2) We established a new assay to precisely measure the degradation rate from the immobile fraction based on FRAP analysis for more than 10 hours, 3) we measured the decay length of the mobile and immobile pools and 4) we added a theory to discuss the effects of growth and non-stationarity.

The reviewer raises here a very important question: what happens if the system is not in quasi-steady-state? Indeed, the publication by Fried and Iber, 2014[17] discusses a situation where growth can extend a gradient with time. This effect occurs in particular in a regime where the turnover of molecules is slow and comparable or slower than the rate of tissue growth. If the ligand molecules turn over faster than the growth rate, a quasi-steady state assumption holds and the gradient is not stretched by growth. If molecules turn over very slowly, they are carried in the growing tissue and the gradient can stretch via growth.

In the new version of the manuscript, we have established a novel assay: long-term FRAP in both small and large discs. The new experiments confirm that the turn-over of molecules occurs with different dynamics in two different pools: a fast mobile pool and a slow immobile pool. Therefore, we have to discuss both pools separately to address whether the steady state assumption is valid. In the new long-term FRAP experiments, the recovery is measured over ten hours, while the standard FRAP experiments that we performed before was only for one hour. With this new long-term FRAP assay, we can estimate robustly the time scales of the fast and slow pools, in particular the time scale of relaxation of the slow immobile pool (NEW Extended Data Figure 11a-c).

Fast pool. The fast pool relaxes within half an hour (Extended Data Figure 1j). This rate is much faster than the growth rate in both early and late discs, where the tissue size doubles in 20h and 25h, respectively (NEW Extended Data Figure 11e-g). As requested by the reviewer, these growth rates were measured again, but now for the precise genetic conditions considered in this report (NEW Extended Data Figure 11e-g).

In the new version of the manuscript, we established another new assay based on FRAP to measure the decay length of the total pool gradient (λ_T), the mobile pool gradient (λ_M) and the immobile pool gradient (λ_I). In this assay, we first measured the decay length of the steady-state gradient before bleaching (λ_T), we bleached GFP-Dpp in the entire target area of the posterior compartment, and then we measured the decay length of the mobile pool (λ_M) after half an hour of recovery (NEW Extended Data Figure 11h,i). Note that, indeed, in half an hour the mobile pool has recovered completely, while the immobile pool is not recovered yet. We found that $\lambda_T \simeq \lambda_M$ (since $\lambda_T \simeq \lambda_M$, then $\lambda_T \simeq \lambda_I$ too). This shows that the fast-relaxing mobile pool quickly forms a gradient at the appropriate decay length (*i.e.* a decay length which scales with the size of the tissue) and that this gradient is in a quasi-steady state (recovery is much faster than growth). Therefore growth has no impact on the decay length of the mobile pool gradient. This validates our parameterization.

Indeed we have estimated the parameter values of Dpp transport based on fast dynamics of the mobile pool which is in steady-state in a short time scale: 1) the fast dynamics of the nanobody internalization assay (30 minutes); 2) the fast dynamics of recovery of the mobile pool in FRAP; 3) the size of the extracellular pool, dominated by rates (k_{on}, k_{off}, D_0, k) which are all much faster than the growth rate and 4) the decay length of the total pool (λ_T). As we consider above, the decay length of the total pool poses a problem, because it contains a fast and a slow component. However, we show above that $\lambda_T \simeq \lambda_M$ and therefore the decay length of the total pool serves as a proxy for the decay length of the mobile pool.

In summary, our parameterization assays are based on the mobile pool, which is in steady-state. This validates the assays, since the dynamics of the assays are much faster than growth. Our validated parameterization

uncovers that the scaling of the mobile pool, which is in quasi steady-state, is mediated by changes of k_{on}, k_{off} and D_0.

Slow pool.

The gradient of the slow, immobile pool has to be discussed separately. The immobile pool relaxation is slower than the growth rate and is therefore not in steady-state. Note however that the immobile pool not being in steady state is incorporated in our model and this does not invalidate our steady state assumption for the mobile pool and the parameter estimates. For instance, when we performed the Bayesian estimation to fit the FRAP data, and in other numerical estimations, we used the dynamic equations (equations S1.1 to S1.5); we did not make any steady-state assumption about the immobile pool, but instead consider the full dynamics.

In order to better understand the properties of the immobile pool and in particular why it scales with system size, we have added a discussion of the theory that takes into account a slow and a fast pool in a growing tissue (new chapter 3.9 in Supplementary material).This analysis reveals (see our new chapter 3.9 in Supplementary material Section 2) that, the immobile pool gradient is not reaching a steady state, but its decay length is set by the decay length of the mobile pool: both the decay lengths of the mobile and immobile pools change in parallel without delay ($\lambda_M \simeq \lambda_I$) and only the amplitude of the immobile pool gradient relaxes slowly with time. Our analysis describes the dynamics of relaxation of the immobile pool in a tissue that grows, assuming that the mobile pool is at quasi steady state so that,

$$C(x,t) = C_0(t)\xi\left(\frac{x}{L(t)}\right) \qquad (1)$$

$$\partial_t C_I + v_x \partial_x C_I + g C_I = \bar{k}_i C - k_2 C_I \qquad (2)$$

where C and C_I are the concentrations of the mobile and immobile pools respectively, v_x is the local cell velocity due to growth, $\bar{k}_i$ is the rate at which molecules are transferred from the mobile to the immobile pool, k_2 is the degradation rate in the immobile pool and g is the growth rate defined by $g = \partial_x v_x(1+\epsilon)$, where ϵ is the tissue growth anisotropy defined by $\epsilon = \partial_y v_y / \partial_x v_x$. $\xi(r)$ is the scaling function which captures the time invariant shape of the gradient. Eq. (2) can be solved to obtain the time dependent expression of C_I given by

$$\tilde{C}_I(r,t) = \bar{k}_i\, e^{-\left(k_2 t + \int_0^t dt'\, g(t)\right)} \int_0^t dt' \left(C_0(t) e^{\left(k_2 t + \int_0^t dt'\, g(t)\right)} \right) \xi(r) \qquad (3)$$

where $r = \frac{x}{L}$. This shows that the amplitude (in red) of the immobile pool relaxes slowly (it depends on time) when the decay length of the mobile pool changes during scaling. In contrast, the decay length of the immobile pool (captured in the scaling function $\xi(r)$, which does not depend on time) follows instantaneously and expands via growth to remain equal to the changing decay length of the mobile pool. This is consistent with our FRAP data described above showing that the decay length of the mobile pool is the same as the decay length of the total pool ($\lambda_T \simeq \lambda_M$) and therefore also to the decay length of the immobile pool. Therefore, the dynamics of the decay length of the total pool scales during growth and reflects the changes of transport parameters which set the scaling decay length of the mobile pool.

In the new version we wrote, when discussing the assays in the main text

> *page 4. The immobile pool of Dpp relaxes in a time scale which is similar to the growth rate of the disc (Extended Data Fig. 11g), raising the issue whether the gradient is in steady state as previously discussed[17]. However note that the assays capture the fast dynamics of˙Dpp gradient formation, which are much faster than the growth rate of the disc. We also found that the decay-length of the total pool of Dpp is not different to the steady state decay length of the mobile pool (Extended Data Figure 11h,i). This justifies the quasi steady-state assumption for the decay length underlying the parameterization. This is discussed in detail in SM 3.9.*

Overall, the authors have spent a lot of thought on the analysis, as reflected by the extended Supplementary Material. The experimental data in this study exploits the most recent experimental technologies and presents an important advance over existing data, even if a careful reanalysis may change the conclusion regarding the scaling mechanism.

Suggested improvements in a revision:

Line 182: “This gradient scaling must be mediated by changes in transport rates.” This statement is not generally true – scaling is possible also for constant transport rates as long as the gradient is not in steady state (Fried & Iber, Nat Commun, 2014). – but only if advection-dilution dominates degradation.

Please check the steady-state assumption with the new data, and re-analyse the data with a pre-steady state gradient, if appropriate.

In the new version of the manuscript, we consider the possibility that the system is not in quasi steady-state. We addressed this in the previous point.

Unlike in the 2011 study, no growth curve appears to be provided. It would be important to provide such a growth rate to be able to solve the model on a growing domain and evaluate scaling; the growth rate is likely lower, given the 10-fold lower Dpp expression in the new system.

In the new version, we measured the growth rate in the new conditions, as explained above (Extended Data Figure 11e-g).

The authors state that their estimate of the extracellular fraction hinges on a large decay length limit – again, how large is the error? Does it explain the inferred differences between young and old discs?

The "large decay length limit" which we used for the extracellular fraction ρ is the limit where the decay length is large which simplifies our analytical expression to be

$$\rho_{\lambda\to\infty} = \left(1 + \frac{k_{on}k(1 + k_i/k_2)}{(k_{on} + k_{\text{off}})(k_r + k_o) + kk_o}\right)^{-1}$$

The full expression is

$$\rho = \frac{k_2(k_0\left(2e^{\frac{1}{\lambda}}k + 2e^{\frac{1}{\lambda}}k_{off} + (1 + e^{\frac{1}{\lambda}})k_{on}\right) + (2e^{\frac{1}{\lambda}}k_{off} + (1 + e^{\frac{1}{\lambda}})k_{on})k_r\right)}{\left(1 + e^{\frac{1}{\lambda}}\right)kk_ik_{on} + k_2(k_o\left(2e^{\frac{1}{\lambda}}k + 2e^{\frac{1}{\lambda}}k_{off} + \left(1 + e^{\frac{1}{\lambda}}\right)k_{on}\right) + 2e^{\frac{1}{\lambda}}k_{off}k_r + \left(1 + e^{\frac{1}{\lambda}}\right)k_{on}(k + k_r))}$$

The following table shows that this approximation introduces errors which are below 2%:

Condition	$\rho_{\lambda\to\infty}/\rho$
L=80µm	0.987
L=144µm	0.987
pent^2 mutant	0.986

therefore the difference in extracellular pool we saw between young and old discs ($\rho = 0.43 \pm 0.02$ *versus* $\rho = 0.13 \pm 0.01$) are not a consequence of considering the "large decay length" limit. We now explain this in the description of the "extracellular pool assay" in the Supplementary Material.

In the model description in Supp 1, the domain and with it the values of v_n are not well explained, i.e. for which n is v_n = 1? This is better explained in the preprint, but it would be good to have all information in this Supp text.

In the Supplementary Material Section 1, in the chapter describing the Steady-State concentration profiles, we mention the values that ν_n takes and under which conditions:

> *"Equations (S1.1) to (S1.5) can be solved at steady state in a piece-wise manner for regions of constant production rate ν_n. We have ν_n= 1 if the nth cell is cell inside the source and 0 if not."*

We have now added a sentence introducing the production rate in the chapter describing the dynamic equations:

> *"The production rate for cell n is given by ν_n and is non-zero only inside the source."*

Pulse-chase experiments: it is not clear from section 2.4 what was measured (where is the data?) and how it was analysed. I may have missed a more thorough explanation elsewhere in the Supp. This study provides a lot of data and details. It would be helpful if you could link to it – or add it if it is indeed missing.

The values of k_2 in the previous version were a rough approximation corresponding to the observations in Entchev 2000, where Dpp was pulsed from the source capitalizing on the thermosensitivity of the gal4 system.

The question of the reviewer about stationarity prompted us to develop a more accurate assay based on long term FRAP, where the recovery of GFP-DPP in the bleached ROI is almost complete. We described this new assay in detail in the Supplementary material of the new version.

FRAP experiments: also, here it is not fully clear how this was analysed.

Supp Section 2.5.2: it is not clear how k_"on" , k_"off" , k_i, D_0 and k_2 were exactly determined – the description is rather vague and no equations are provided.

In the new version we have strengthen the description of the method in the chapter 2.5.2 in the Supplementary Material Section 2. We made explicit references to the dynamic equations and to precise figure items to clarify the description. Equations for the determination of the intervals for all the parameter values except k_{on}, k_{off} and D_0 are given in chapters 2.1 to 2.5 in the Supplementary Material Section 2. There are no explicit equations for k_{on}, k_{off} and D_0. The determination of these parameters is based on an approach applying Monte Carlo sampling and Approximate Bayesian Computation to the FRAP data. We outline our method in detail in Supplementary Material Section 2.5.2 and summarize this here briefly.

Dpp transport depends on eight parameters. Three of them (endocytosis, recycling, clearance and immobilization rate) are determined by means of the nanobody uptake assay (explained in Supplementary Material Section 2.2). A fourth rate, the degradation rate of the immobile pool, was determined by measuring the slow relaxation rate in a long-term FRAP experiment. The last three, the extracellular diffusion and the binding/unbinding rate were determined by means of the Monte Carlo sampling and Approximate Bayesian Computation (Supplementary Material Section 2.5.2).

In summary, the rationale of the Bayesian approach is that these remaining parameters are constrained by i) the FRAP data, ii) the experimental determination of the decay length λ in steady state, iii) the experimental determination of the extracellular fraction ρ and iv) the long term FRAP assay. To narrow the ranges of values for these parameters, we considered a large number ($3 \cdot 10^7$) of randomly chosen sets of parameter values (Extended Data Fig. 8g). For each set of values, we calculated the expected relaxation in the FRAP experiment using the dynamic equations for the five pools of Dpp (equations S1.1 to S1.5). We then compared those solutions to the experimental relaxation measured in the FRAP assay and selected those with high quality of fit. This determines sets of possible values consistent with all five assays and, in particular, constrains the values for the extracellular diffusion coefficient, the binding and unbinding rates.

All analysis scripts and data files should be made available to the reader. Even though the authors provide extensive Supplementary Material, it often would be easier to see what was done if one had access to the code; this would directly address the above unclear points. Also, others can build more easily on the work.

We will make this available as requested by the reviewer.

The paper heavily rests on Reference 4, which is listed as published in Phys Rev E in 2019. However, I could only find a preprint https://arxiv.org/abs/1909.13280.

We now have corrected this in the reference list.

Small Points:

Line 45: Please explain how it follows that photoconverted molecules are only in the endocytic, but not in the secretory pathway. If there is not enough space, please refer to the Supplementary Material.

Because of the way the experiment is done, it can be excluded that the origin of the photoconverted molecules that spread is exclusively secretory. This is because, in the experiment, GBP-Dendra2 nanobodies are first internalized and the extracellular pool of nanobodies is eliminated by an acid wash which is efficient (see NEW Extended Data Figures 5e,f and 8a-d). After acid wash, only endocytosed GBP-Dendra2 molecules remain in the tissue and it is only then that Dendra2 is photoconverted: this internalized pool is the only source of fluorescent molecules that can spread in the tissue. Therefore the origin of GBP-Dendra2 cannot be exclusively secretory.

It is however possible in principle that part of the internalized pool traffics through a secretory compartment. In this scenario, the route for Dpp recycling would involve the secretory pathway (for which there are some examples in the literature[2,3]), which is indeed a potential mechanism of recycling both in the source domain and in the target. The alternative to this route involving the secretory pathway involves recycling endosomes[4].

In the revised version, we also eliminated the strict exclusion of the secretory pathway as a possibility:

> *page 2. Because the photoconverted molecules were only in the endocytic pathway, ~~and not in the secretory pathway~~ this result indicates*

This scenario of endocytic/secretory mixage would only be an issue in the case that the two pathways can only mix in the source domain, but cannot mix in the target tissue. In this hypothetical case, the photoconversion of internalized GBP-Dendra2 in the source domain does not exclude that, in the target cells, endocytosed molecules may not be re-externalized again. To exclude this possibility, and show that endocytosed molecules are re-externalized from target cells, we performed a new series of experiments where we incubated with GBP-Dendra2 and performed acid wash as before, followed now by photoconversion in a ROI in the target tissue, which does not overlap with the source domain. The NEW Figure 1b shows that endocytosed GBP-Dendra2-Dpp molecules do spread from target cells and can form a gradient. In the main text of the revised version of the manuscript we added:

> *page 2. Photoconversion of this intracellular GFP-Dpp-bound GBP-Dendra2 in a region of interest (ROI; whether in the source (Fig. 1a) or target tissue (Fig. 1b)) allows us to address whether these internalized molecules can move to neighbor cells and form a gradient there (for details and controls of this assays, see SM 2.8; Extended Data Figures 5e-g, 8a-f).*

and

> *page 2. Because photoconversion in the target territory also lead to formation of a gradient (Fig. 1b), this shows that Dpp molecules can be re-exocytosed also from the target cells.*

Figure 1a: add scale bar

We added the scale bar.

Figure 1b: where is zero? Is this the red line in a?

Zero corresponds to the source boundary and the red line in a. This is now explained in the legend.

How does the amplitude evolve in the part that was photo-converted?

We have measured the fluorescence of photoconverted GBP-Dendra2* at the source boundary as a function of time and compared our experimental curve with the curve generated by performing the photoconversion experiment numerically using the dynamic equations for the five pool (S1.1-S1.5) (see NEW Extended Data Figure 5h). The dynamics of the experiment and the numerical simulation for the fitted parameters are very similar, providing an independent validation of our parametrization procedure.

Figure 2f,g: Why was the pent mutant not measured also at a short length?

The parameterization of the eGFP-Dpp in small *pentagone* mutant discs is challenging, because the eGFP-Dpp signal intensity is particularly low in this condition. Nevertheless, following the suggestion of the reviewer, in the revised version of the manuscript, we have performed the parameterization of the *pentagone* mutant in the small discs. In wildtype small discs (l=80 um) the contribution of the recycling and the unbound module were the same. As expected, in *pent*^{2} mutant discs of the same size, the unbound module predominates over the recycling module, as we have already observed in the large *pent*^{2} mutant discs (see NEW Extended Data Figure 15e,f). We refer to it in the main text.

Figure 2f,g, 3a: it is maybe misleading to present this type of information as bar plot – a box plot would be more appropriate.

We have now presented the data in Figure 2g as a box plot. The nanobody uptake parameters in the chart in the Figure 2f are determined from a simultaneous fit (with its confidences) of the entire dataset and not as averages from individual fits (see Supplementary Material Section 2 chapter 2.2.3) it would therefore not be evident how to represent this data as a box plot. In Figure 3a, in order to put all of the parameters in one chart, we preferred to represent the numbers in the logarithmic scale, for which box plots are not convenient. To be transparent about the spread of the data, we represented all the individual data points behind Figure 3a in pairwise plots in Figures 3c,d and Extended Data Figures 2a, 5i and 6e).

The work is spread over five different files (main file, Extended Data, Supp 1, Supp 2, separate manuscript), which makes it hard to follow – but maybe there is no better way.

Referee #3 (Remarks to the Author):

Gonzalez, Jülicher and colleagues combine experimental data and mathematical modeling to propose that scaling of the Dpp morphogen gradient in the growing developing wing disk of Drosophila is achieved by increasing the contribution of the internalized Dpp molecules to the extracellular pool of Dpp. The topic is highly relevant for the developmental biology community and different models have been proposed to address this beautiful question. By combining a good number of newly generated tools with nicely developed and very imaginative experimental settings, authors provide convincing evidence that the pool of Dpp endocytosed is being recycled and sent to the extracellular pool again, that this recycling contributes to gradient formation and that Dpp endocytosis is independent of the Tkv receptor. I have no major issue with this part and as a whole is a very interesting observation.

We thank the reviewer for their appreciation of our work.

Whether changes in Dpp recycling functionally contributes to scaling and whether pentagone is indeed an important player in morphogen scaling are two relevant questions that are not convincingly addressed in this ms, though.

In this context, a recent report of the lab of Arthur Lander in Dev Cell seriously questions the long-range activity of pentagone and rules out a potential role of this gene in Dpp scaling. Indeed, it is proposed in this paper (Zhu et al 2020) that a feedback negative loop between Dpp and its receptors/coreceptors appears to be the one mediating morphogen scaling and it is shown that pentagone activity is rather local and restricted to the lateral regions of the Dpp gradient.

In order to address the concerns of the reviewer in detail, we would like to discuss the following points:

1) The authors of the Zhu paper do not themselves claim that Pentagone does not play a role in scaling[18]. Indeed, the relevant paper does confirm itself the scaling phenotype of the Dpp signaling gradient in *pentagone* mutant discs which has also been previously described by various reports from different labs[19,20]. This is consistent with the effect of *pentagone* knockout on Dpp ligand/signaling scaling reported in our manuscript (Extended Data Figure 6a,c);

2) The role of HSPGs as receptors/coreceptors in Dpp gradient formation (and scaling) proposed by the Zhu et al paper is consistent with the observations in our manuscript and other reports[19,21];

3) The claim in the Zhu et al paper that Pentagone activity is cell autonomous and only operates in a short-range is not confirmed by our observations and, in turn, our observations on the non-cell autonomy of *pentagone* are completely consistent with previous reports from several other labs, including those from the labs that discovered Pentagone (Affolter, Pyrowolakis)[13,22].

1) It is important to emphasize here, that the paper mentioned by the reviewer[18] does not claim Pentagone is not essential for scaling. The authors in Zhu *et al* write:

> "*As discs grow, Dpp gradients expand, keeping relative boundary positions approximately stationary. Such scaling fails in mutants for Pentagone (pent), a gene repressed by Dpp that encodes a diffusible protein that expands Dpp gradients.*"

Moreover, similarly to other reports from the Affolter and Shilo labs (Hamaratoglu *et al.*[20], Fig. 7, Ben-Zvi *et al.*[19], Fig. 2), Figure 5C in Zhu et al shows that, in pentagone mutants, the Dpp signaling gradient does not scale: the decay length of the P-Mad gradient does not increase with the increasing size of the wing disc.

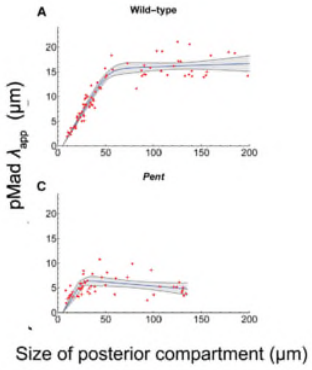


Figure 5 in the Zhu *et al*. paper. Feedback Regulation of tkv and dally Is Required for Scaling

(A–H) pMad (…) decay lengths in the wing disc posterior compartment were measured and fit (…). Envelopes around curves represent 95% confidence intervals for the fits. Genotypes were (A) wild type; (B) +/pent2; (C) pent2/pent2

Note that in previous reports (the two reports above from Affolter and Shilo and Zhu *et al.*, 2020), the *pentagone* mutant scaling phenotype has only been observed for signaling readouts (P-MAD) and data for Pentagone-dependent scaling for the Dpp ligand itself has never been shown before. In this manuscript, we did confirm the pentagone-dependent scaling of the signaling gradient (Extended Data Figure 6c). Furthermore we show an unambiguous scaling phenotype of the Dpp ligand gradient itself in *pent*^{2} mutants using eGFP-Dpp^{LOP} (Figure 2a; Extended Data Figure 6a).

In summary, the importance of Pentagone for Dpp gradient scaling is not questioned in Zhu *et al.* Instead, Zhu *et al.* questions the mechanism by which Pentagone affects scaling, they question the Expansion-Repression (ER) model:

> "*Here, we re-examine the role of Pent in the Drosophila wing disc, focusing on an important requirement of the ER model, namely that the expander spread in a nearly uniform manner over the morphogen field.*"

The authors' claim that Pentagone is cell autonomous prompts them to exclude the ER model and propose an alternative model for scaling of the Dpp signaling gradient by Pentagone, in which Pentagone functions in concert with co-receptors (Dally/Dlp). Therefore, Zhu et al do not question whether Pentagone is essential for scaling, but the mechanism by which it functions.

2) The role of HSPGs as receptors/coreceptors in Dpp gradient scaling proposed by Zhu et al is consistent with the observations described in our manuscript: we show that, like in *pentagone* mutants, *dally* mutant discs present a scaling phenotype both in Dpp signaling and Dpp ligand gradient (Extended Data Figure 6b,d). Indeed in our report, we show that Dally is as important as Pentagone in controling scaling and that they both share the mechanism by which they mediate scaling: by modulating the diffusion, binding and unbinding rates D_0, k_{on} and k_{off} (Extended Data Figure 6e,f). We speculate on how HSPGs together with Pentagone could affect D_0, k_{on} and k_{off} in the Supplementary Material Section 2 chapter 4.5. Note that feedback loops of Dpp with receptors/coreceptors (Dally/Dlp), proposed in Zhu *et al.*, have not been studied in our work and are still compatible with our model.

3) Important claims in Zhu *et al.*, are that Pentagone activity is cell-autonomous and that Pentagone spreads only on a short range (a few cells away from the expression domain). Two observations are invoked to support these claims: **a)** the phenotype of Pentagone overexpression in clones and b) the range of spread of GFP-Pentagone from its source.

a) The authors claim that overexpression of Pentagone in flip-out clones reduces Dpp signaling in a cell-autonomous manner as seen by P-Mad staining (Figure 3A-B,A'-B').

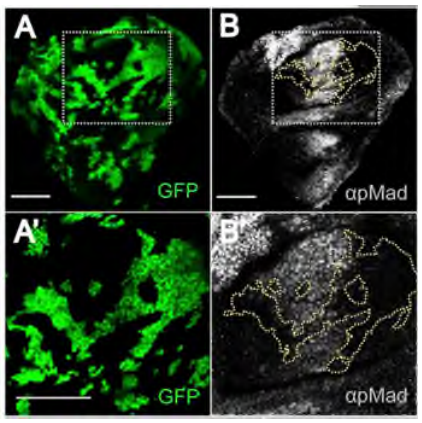


Figure 3 in the Zhu *et al.* paper. Pent Overexpression Has Only Local Effects

(A–E0) Third instar wing disc with clones of cells overexpressing pent (GFP, green; A, A') stained with anti-pMad (gray; B, B')

It is worth noting that the data shown corresponds to one disc, but they only had five discs in their sample size. Otherwise, there are two issues with this observation: i) in their hands, overexpression of Pentagone causes loss of Dpp signaling and ii) that phenotype is cell-autonomous.

It is well established that loss of function *pentagone* mutants reduce Dpp signaling, causing the range of the gradient to shrink, as observed by different labs[13,19,20,22], including the observations in Zhu *et al.*[18]. Therefore, it is very surprising that Zhu et al observe a loss of function for Dpp signaling upon Pentagone overexpression in clones (Fig 3A,B). Indeed, other labs have published that overexpression of Pentagone causes a gain of function (*i.e.* expansion) Dpp signaling phenotype (Vuilleumier *et al.*[22], Fig. S7, Ben-Zvi *et al.*[19], Fig. 1,3). Surprisingly, Zhu *et al.* also acknowledge the gain of function phenotype of Pentagone overexpression:

> *p.725 "Loss of pent dramatically shrinks the Dpp signaling gradient, and pent overexpression expands it (Vuilleumier et al., 2010)"*

These experiments were performed in conditions in which Pentagone was overexpressed in all the cells within a compartment, while Zhu et al overexpress Pentagone in clones. We and others (Pyrowolakis lab, personal communication) have however performed an identical experiment in clones in which both the genetics and molecular reagents used as well as the temperature/time details of the protocol are the same as in Zhu *et al.* In the Figure below (from a manuscript on pentagone which is currently under consideration in other journal), we observed that i) in contrast to Zhu et al, PMad levels are not reduced in the clones, instead the presence of Pentagone overexpressing clones causes the expansion of the range (decay length) of the PMad gradient in the whole compartment. This result indicates that the overexpression of Pentagone in clones causes a non-cell autonomous expansion of the gradient. It is also important to remark that overexpression of Pentagone does not cause a loss of PMad staining.

[REDACTED]

[REDACTED]

Note that the long-range activity of Pentagone is confirmed by the fact that overexpansion of the Dpp gradient upon overexpression of Pentagone in one compartment (anterior or posterior) can be observed in both compartments (anterior and posterior), including the compartment where Pentagone has not been overexpressed. This was shown in previous report such as Vuilleumier *et al.*[22] (their Fig. S7) and Ben-Zvi *et al.*[19] (their Fig. 1,3).

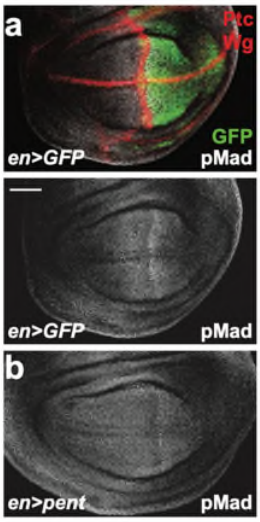


Figure S7 in Vuilleumier *et al*.

Pent acts in a feedback loop in BMP morphogen gradient formation. (a-e) pMad staining (grey in a, b), (…) in wings expressing *GFP* (a) or *pent* (b) in the P compartment. Note the strong smear of pMad in the A compartment.

Surprisingly in Zhu et al, the authors acknowledge this result which strongly supports non-cell autonomy far from the source where pentagone is overexpressed. Zhu et al uses this result to reject the possibility that the A/P boundary represents a barrier to pentagone function. They write:

> *p. 725 "(…) given the ease with which overexpressed Pent in the (entire) posterior compartment can produce phenotypes in the anterior*"

Together, these observations in different reports confirm that Pentagone activity is not cell-autonomous and long-range.

b) To support their claim that pentagone acts short range and cell autonomously Zhu et al looked at overexpressed pentagone-GFP driven by Hedgehog-gal4, Apterus-gal4 and Cut-gal4 and argued that the spread from the source was short (their Figure 2A,A').

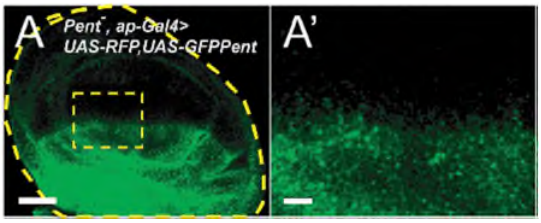
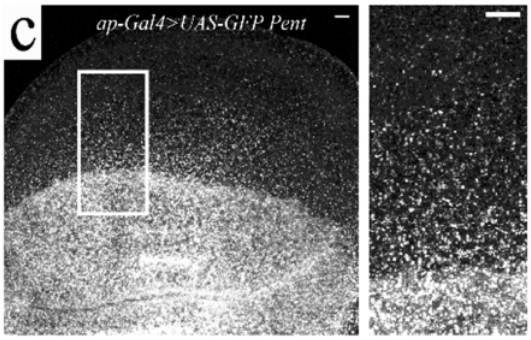


Figure 2A,A' in Zhu *et al*. The Apparent Decay Length (lapp) of GFP-Pent Is Short

(A–C) UAS-GFP-Pent expression was driven by ap-Gal4 (dorsal domain, A–A')

Using the same reagents, we overexpressed Pentagone-GFP driven by Apterus-gal4 and observed that Pentagone is making a gradient with a substantially longer decay length than reported in Zhu *et al.*, 2020. (see NEW Extended Data Figure 15c,d).



Extended Data Figure 15c in our revised manuscript. UAS-GFP-Pentagone expression driven by ap-Gal4 (corresponding conditions as in Zhu et al 2020). Higher magnification of the fluorescent signal of the area boxed in the image to the left is shown to the right. Scale bars, 10 µm.

In the revised version of our manuscript, we have now added a discussion on the observations in Zhu *et al.* (Supplementary Material Section 2 chapter 4.1).

The contribution of Dpp recycling to Dpp scaling appears to be rather correlative and its proposed causality is not demonstrated.

In our manuscript, we challenge the role of recycling for gradient formation by proposing alternative, independent assays (beyond those used for parameterization), that specifically address the contribution of recycled molecules to the formation of the gradient (photoconversion and iFRAP experiments; Figures 1a,b; 4c-g). In these experiments we use nanobodies which are internalized together with Dpp into cells, then we remove the extracellular nanobodies bound to GFP-Dpp by means of an acid wash and interrogate those internalized molecules (by means or FRAP or photoconversion) for their capability to be re-secreted, to detach from the membrane, to diffuse and to reappear in neighbour cells. They can do all that, indicating that their recycling can generate a gradient. These experiments establish, beyond mere correlation, that recycling contributes to gradient formation.

In addition, to further investigate the role of recycling in gradient formation, in the revised version of the manuscript, we look at the effect on the Dpp gradient of downregulation of recycling. In order to downregulate recycling, we have downregulated the expression of the recycling Rab proteins: Rab11 and Rab4. We perform RNAi experiments in which we express validated dsRNAs for Rab11[23] or Rab4[24] in the posterior compartment of the wing disc in a temperature dependent manner for 4 days before dissection. We have observed a decrease of the decay length of eGFP-Dpp^{LOP} gradient (see NEW Extended Data Figure 12a-c), indicating that recycling plays an important role in gradient formation. We have added the description of this experiment to the Supplementary Material Section 2.

> *We show that large discs are in a combined transport regime, where extracellular diffusivity, recycling and unbinding are all important for the formation of the gradient. Consistently, the boosts B_{k_r}, $B_{k_{off}}$ and B_{D_0} are important (Extended Data Figure 3c). Consistent with a large B_{k_r}, the decay length of the gradient is shortened upon partial silencing of recycling by RNAi through expression of dsRNA for the recycling Rabs (Rab4 and Rab11) in the posterior target cells (Extended Data Figure 12a-c).*

Authors do provide evidence that Tkv is not mediating Dpp endocytosis. What about Dally and Dlp?

In the previous version of the manuscript, we showed that, by itself, Dally mutations do not affect the endocytosis rate (Extended Data Figure 6e,f). In order to still address whether Dally together with Dlp have an effect in endocytosis, we abolished both their function by cleaving the two proteins away from their GPI-anchor at the plasma membrane through incubation with the PI-PLC enzyme. We then performed a nanobody internalization assay for GFP-Dpp. Under these conditions, we observed a considerable reduction of endocytosis and recycling rates (see NEW Extended Data Figure 14e,f). These results indicate that together, Dally and Dlp might play a role in endocytic trafficking.

Taken all these issues into consideration, I am not convinced that this ms is a serious candidate for Nature.

Referee #4 (Remarks to the Author):

Gradients of diffusible morphogens provide positional information in growing tissues, suggesting that gradients scale with the size of the territory. Deciphering the molecular and cellular basis of scaling is an important issue. Wolpert (1969) first realized that a gradient formed by linear diffusion from a source to a sink scales naturally as the sink moves away from the source. This model, however, requires a local sink and an otherwise stable morphogen. Building upon this idealistic view, Barkai (2010) proposed a model of automatic gradient scaling in which morphogens repress the expression of a diffusible expander that negatively feeds back on morphogen production. In the context of the Dpp gradient in the developing fly wing, a textbook model in Developmental Biology, scaling was proposed to involve the secreted protein Pentagone as an expander of the Dpp gradient. Here, Gonzalez-Gaitan and colleagues study the scaling of the Dpp gradient and propose a related model. In this

new model, Pentagone mediates gradient expansion by increasing the rates of Dpp binding to its putative internalization receptor. This is a potentially interesting idea since the secreted protein Pentagone was recently shown to diffuse only minimally (Zhu et al., 2020, Dev Cell, 53, 724-39).

The key findings/claims of this paper (abstract) are that:

- scaling (increase in decay length) is achieved by increased recycling of Dpp

- increased recycling results from changes in the binding rates of Dpp to its internalization receptor

- these binding rates are modulated by Pentagone

Many data (reviewed below) are provided to support these statements. However, to my best understanding, these data provide only indirect evidence in support of the three claims above. Indeed, the rate of Dpp recycling was not measured. Thus, whether this rate changes over time and is modulated by Pentagone remain to be determined.

First, note that the recycling rate k_r has indeed been measured by the nanobody uptake assay (Fig. 2b-f; Extended Data Figure 1f,g). Cells uptake a nanobody against GFP in GFP-Dpp expressing discs with dynamics which corresponds to an exponential which relaxes to a linear regime. These dynamics are captured by three parameters: p, the exponential relaxation time; B, related to the saturation of this exponential and A, the slope of the linear dynamics. We measured experimentally p, B and A by fiting data to an equation which captures the exponential relaxation into a linear regime. We then showed that the recycling rate is a function of p, B and A by $k_r = p^2B/(pB + A)$. All this was explained in detail in Supplementary Material Section 2.2.

Using this approach, we showed that the recycling rate does indeed not change in small, large and *pentagone* and *dally* mutant discs. Instead what changes is the contribution of the RECYCLING MODULE to the Dpp gradient, which happen by changes in D_0, $\mathrm{k_{on}}$ and $\mathrm{k_{off}}$.

The recycling rate (k_r) and the recycling module are not the same thing. In our analysis, we introduced the concept of modules that capture the contribution of different trafficking processes to the formation of the gradient. We showed that in the three conditions (small discs, *pent* and *dally* mutants), the contribution of the recycling **MODULE** to the decay length is smaller than in large discs at the expense of the unbound **MODULE** (see Fig. 3b and Extended Data Figure 6f), and this happens without an effect on the internalization or recycling rates, but by changing the extracellular diffusion (D_0) and the binding/unbinding rates ($\mathrm{k_{on}}$, $\mathrm{k_{off}}$) to the receptor which internalizes Dpp (see Fig. 3a). The concept of modules and the difference between modules and transport rates is further

described in chapter (3.1) in the Supplementary Material Section 2, where we explain in particular the difference between the recycling module and the recycling rate. Moreover, to clarify this point, in the new version of the main text of the manuscript we wrote:

> *p.3. Note that the recycling module is not equivalent to the recycling rate, as it depends on all the transport rates, except k_2. Importantly, extracellular diffusivity and the binding rate are essential for the three modules described above.*

In fact, internalization and recycling rates are not affected in these conditions. Only D_0, $\mathrm{k_{on}}$ and $\mathrm{k_{off}}$ rates change considerably. In other words, variations in the relative contribution of the recycling module in our analysis is not due to changes in the recycling rate (or any other intracellular rate, including the internalization rate). Instead, the relative contribution of the recycling module increases in larger discs due to changes in the extracellular rates (D_0, $\mathrm{k_{on}}$ and $\mathrm{k_{off}}$) that make intracellular molecules available to return to the extracellular unbound pool and diffuse to spread and contribute to a longer decay length of the gradient.

Additionally, the rates of Dpp binding to its receptor were not directly measured. Therefore, whether these binding rates change over time and depend on Pentagone function remain to be established.

We did estimate the binding rates by means of an approach based on Monte Carlo sampling and Approximate Bayesian Computation. We outline our method in detail in Supplementary Material Section 2.5.2 and summarize this here briefly.

Dpp transport depends on eight parameters. Three of them (endocytosis, recycling, clearance and immobilization rate) are determined by means of the nanobody uptake assay, similar to what was explained above for the recycling rate (explained in Supplementary Material Section 2.2). A fourth rate, the degradation rate of the immobile pool, was determined by measuring the slow relaxation rate in a long-term FRAP experiment. The last three, the extracellular diffusion and the binding/unbinding rate were determined by means of the Monte Carlo sampling and Approximate Bayesian Computation (Supplementary Material Section 2.5.2).

In summary, the rationale of the Bayesian approach is that these remaining parameters are constrained by i) the FRAP data, ii) the experimental determination of the decay length λ in steady state, iii) the experimental determination of the extracellular fraction ρ and iv) the long term FRAP assay. To narrow the ranges of values for these parameters, we considered a large number ($3 \cdot 10^7$) of randomly chosen sets of parameter values (Extended Data Fig. 8g). For each set of values, we calculated the expected relaxation in the FRAP experiment using the dynamic equations for the five pools of Dpp (equations S1.1 to S1.5). We then compared those solutions to the experimental relaxation measured in the FRAP assay and selected those with high quality of fit. This determines sets of possible values consistent with all five assays and, in particular, constrains the values for the extracellular diffusion coefficient, the binding and unbinding rates.

Of note, the authors propose the existence of a Dpp internalization receptor that is distinct from the signaling receptor Tkv. Because the molecular identity of this putative internalization receptor is not known, analysis of the binding rates cannot be performed.

Because we show that Tkv is not the receptor that internalizes Dpp, it is not relevant to measure the binding rates of Dpp for Tkv with techniques such as plasmon resonance for example. Instead, we developed a novel method based on a Bayesian approach (see above), that allows us to narrow down the range of plausible values for the binding and unbinding rates, consistent with the dynamics of our multiple assays. We do not measure the binding/unbinding of Dpp to a particular known receptor, but the effective rates of binding/unbinding which lead to internalization into the cell.

In this study, the authors first tested whether endocytosed Dpp can contribute to gradient formation. To do so, they used two recycling assays, one based on photoconversion (Fig 1a) and another based on FRAP (Fig 4c,d). Both assays probe the secretion of internalized of GFP-tagged Dpp, i.e. recycling, and measure its diffusion in the target tissue (Fig 1a and 4c,d). The results indicate that Dpp is internalized and recycled and therefore provide support that recycled Dpp contribute to the Dpp gradient (but see points 2, 3 and 6 below). However, the relative contribution of recycled Dpp to the pool of diffusive Dpp that forms the gradient remains to be determined. In other words, these data do not exclude the possibility that the Dpp gradient forms mostly by diffusion of Dpp produced at the source.

We agree with this reviewer that the Photoconversion and iFRAP assays uncover the existence of recycling of Dpp molecules back to the plasma membrane, their subsequent detachment from the plasma membrane and their diffusion to reappear in neighbor cells, but these assays do not provide information about what is the relative contribution of the recycled molecules to the formation of the gradient. To address what is the relative contribution of the recycled molecules to gradient formation, we developed the concept of MODULES which provide information

about what percentage of λ^2 (the square decay length) is generated by the recycled molecules (recycling module) or other modules (unbound, bound, transcytosis modules). The relative contribution of different modules is different in different conditions (small, big discs, *pentagone* or *dally* mutants) in which the relative importance of the recycled molecules is different.

In any case, in the revised manuscript we change our formulation to avoid this misunderstanding:

> *page 2. We first developed an assay to address whether internalized Dpp molecules can contribute to formation of the gradient.*

Next, the authors proposed a model for the transport of Dpp which involves five pools of Dpp and eight transport steps (Fig 1c). This model considers both diffusive and cell-based transport mechanism. Of note, the model considers Dpp transport by transcytosis but not via cytonemes (i.e. cell size, a, is constant). The model describes the time evolution of Dpp accumulation in the tissue in a set of equations. A decay length of the gradient at steady state can be obtained from these equations. Its value depends on eight independent parameters (transport rate values listed in Extended Table 1). Because most of these transport rate values are not experimentally accessible, the authors developed transport assays that were meant to depend on only a subset of these transport rates (Fig 2; see points 3-6 below for detailed comments on these assays).

This approach aimed at constraining the parameter values and at narrowing them down to a limited range of values compatible with the proposed model. In a sense, the experimental data are not meant to challenge the model but to make it work.

We performed here a minimal description of phenomena that we know happen in this tissue: extracellular diffusion, binding unbinding, endocytosis, degradation, etc. These phenomena do not correspond one to one to molecular events (e.g. extracellular diffusion incorporates not just the Brownian motion of molecules, but also their binding and retention by the extracellular matrix (see Supplementary Material Section 2 chapter 4.5.2), the endocytosis rate could correspond to the integration of different endocytic pathways with potentially different internalization rates, etc.). Therefore, our rates are effective rates that capture the extracellular mobility, internalization etc. In this sense the approach is an equanimous, minimal description of the system and not a committed model. By itself this description can capture a wide variety of models and indeed, it was able to reveal different regimes, including the extreme extracellular diffusion model (in *pentagone* mutants), a recycling model (in large discs) (Extended Data Figure 4), and a combination of them (in small discs) where different modules prevail.

The experiments are not designed to prove that the 5 pools and 8 rates exist. These effective pools/rates do exist, but their size and dynamics can be of different magnitudes. The transport of Dpp with 5 pools and 8 rates is not a model, it is a physical description. Therefore, we are not "narrowing them down to a limited range of values compatible with the proposed model". We narrow down the trafficking parameters to a limited range of values compatible with quantitative observations (our 5 assays). We use this framework to parameterize and assess the importance of different phenomena in the formation of the gradient. The model emerges from this parameterization. We do not parameterize to make it compatible with a model. Therefore we disagree that "the experimental data are not meant to challenge the model but to make it work".

In the new version of the manuscript, we now emphasize that we do not narrow parameters down to be compatible with a proposed model. Instead we use an unbiased phenomenological description that takes into account trafficking processes by sets of trafficking modules. We can then compare our results to behaviors found in different models (see Extended Data Figure 4, Supplementary Material Section 2 chapter 4.2.2-4.2.3), which are extreme regimes of Dpp transport still captured by our description:

> *page 5. "To characterize the Dpp transport regime in large discs, we estimate the contributions of four transport modules from equation 1 using the measured rate values. A particular regime is characterized by the prevalence of one or several transport modules over others. Examples of regimes include scenarios where either the unbound module, the transcytosis module or the recycling module prevail (SM 4.2.2). Scenarios with mixed modules are also possible. The relative contribution of a module depends in turn on the actual values of the transport rates (see above). It is worth noting that this approach is equanimous: all possible regimes can be described by the same description of transport (Fig. 1b) and the actual regime emerges from the set of rate values determined experimentally in our approach (Fig. 3a). Therefore, a regime is not a consequence of particular modeling assumptions, but emerges from parameter values determined experimentally."*

Briefly, these transport assays were performed at two time points (l=80 and l=144; see point 4 below) and the results obtained were used to generate a limited sets of compatible parameters values (Fig 3a). Using these parameters, the authors examined the relative contribution of the four modules (Fig 1e) and concluded that

transcytosis of Dpp does not contribute to gradient formation (see point 7 below), that diffusion of unbound Dpp is important at early but not late stages,

Extracellular diffusion is important for both the unbound module and the recycling module. This can be seen by looking at the equations for the recycling module (Fig. 1e,f) where D_0 is present. Indeed, one has to distinguish between pools and modules. See the detailed description and the meaning of the concept of modules in the Supplementary Material Section 2, chapter 3.1. In late discs, the unbound module is not important for gradient formation, but the diffusion of the unbound pool is essential. This can be seen for example in Extended Data Figure 3c: in late discs, where the gradient is formed almost exclusively by the recycling module, the D_0 boost is even more important than the k_r boost (for the concept of "boost" see chapter 3.5 in the Supplementary Material Section 2).

To clarify this, we wrote in the main text of the revised version of the manuscript:

> *page 3. (…) The third term ("recycling module") considers the endosomal molecules, which can be recycled to the plasma membrane, where they can again unbind to contribute to the diffusing unbound pool. This process is counteracted by endosomal degradation or immobilization. Note that the recycling module is not equivalent to the recycling rate, as it depends on all the transport rates, except k_2. Importantly, extracellular diffusivity and the binding rate are essential for the three modules described above (…)*

We also added some clarification in the abstract:

> *page 1. (…) to expand the gradient, endocytosed molecules are re-exocytosed to spread extracellularly. To regulate the contribution of endocytosed Dpp to the spreading extracellular pool during tissue growth, it is the Dpp binding rates that are progressively modulated by the extracellular factor Pentagone, driving scaling.*

and that recycling of Dpp plays an increasingly important role over time (Fig 3b). Mechanistically, they propose that the binding rates of Dpp to its internalization receptor vary over time.

These conclusions critically depend on the precision and quality of the data from transport assays which were used to estimate Dpp transport rates (see points 2-6).

We agree with the reviewer that the confidence and precision of each independent assay is important for the robustness of our final conclusions. As such, we have made a systematic effort to establish how the confidence associated with each assay is reflected on our final conclusions.

First, we have incorporated the confidence intervals of each independent assay to our final fitting routine (Supplementary Material Section 2 chapter 2.5.2). Thus, the impact of small variations around the optimal fitted values for λ, k_N, k_o and k_r are considered in our fitting routine. Second, we have investigated how the constraints imposed by the quality of the fit to the FRAP assay affects our conclusions. We have specifically explored how our conclusions change if the quality of the fit (reflected by the R^2 value) is relaxed, and found that our conclusions are robust to changes in the threshold chosen (Extended Data Figure 10a and Supplementary Material Section 2 chapter 3.7). Third, we have explored how errors in the measured value of the extracellular fraction would potentially change our conclusions. We found that our conclusions were robust to a 250% change in the value of the extracellular fraction in large discs, 125% change in small discs and 165% change in *pent*^{2} mutant discs (Extended Data Figure 10b and Supplementary Material Section 2 chapter 3.7). All these were well above the confidence intervals of our measured values for the extracellular fraction indicating that our conclusions are robust to the relatively small fluctuations found in our measurements. We have now modified Supplementary Material Section 2 chapter 3.7 in order to clarify the robustness of our parameter estimation method.

They also depend on assumptions of the model (see point 1 below). Last, since endocytosis and recycling appear to contribute to the Dpp gradient, the authors studied the internalization of GFP-Dpp and confirmed that the Dpp signaling receptor, Tkv, is not required for the internalization of Dpp, suggesting that a Dpp internalization receptor must exist (Fig 4a).

Finally, the authors studied the GFP-Dpp gradient and transport mechanisms in a mutant condition that is known to disrupt scaling. Previous studies showed that the secreted protein Pentagone is required for the scaling of Dpp signaling activity. Here, detailed analysis of pentagone mutant discs indicated that the Dpp gradient does not properly scale in the absence of pentagone (Fig 2a) and that Dpp transport appeared to primarily involved diffusion (Fig 3b). Thus, the authors propose that scaling relies on a time-dependent increase in the concentration of Pentagone, which would lead to a progressive increase in the rate of Dpp recycling.

One obvious prediction of this model, and of the original ER model as noted by the authors in p.6 l.229, is that Pentagone levels increase over time and that Pentagone levels correlate with Dpp recycling. This prediction was

not examined here. Thus, while it is known that overexpression of Pentagone results in an overexpansion of the Dpp signaling gradient (ref 22), whether this effect is due to changes in Dpp recycling remains to be tested.

In the revised version, using lacZ expressed by the pentagone promoter (reagent from Vuilleumier *et al.*[22]), we show that the transcription levels of Pentagone increase during wing discs growth and are proportional to wing size (see NEW Extended Data Figure 15a,b). This is indeed what is expected in the ER model.

The parameterization of the eGFP-Dpp in small discs overexpressing Pentagone is challenging, because the eGFP-Dpp signal intensity is low and surprisingly variable in this condition. Nevertheless, following the suggestion of the reviewer, we have performed parameterization in small discs overexpressing Pentagone. We observe that in this condition, compared to small control discs, the ratio k_r/k_o (which captures the sorting between recycling *versus* degradation) is increased (14 *vs* 23). This could explain the overexpansion of the eGFP-Dpp^{LOP} gradient that we observe in this condition (the decay length $\lambda = 20.9 \pm 1.8\ \mu m$ *versus* $\lambda = 15.9 \pm 0.54\ \mu m$ in control small discs). It is worth noting that these changes in the ratio k_r/k_o are unlikely to be physiological, since in *pentagone* mutants we did not observe a change in the k_r and k_o rates (see Fig. 3a).

In sum, this study addresses a long-standing and debated issue. In order to help the field move forward, robust and precise data are needed if they are to be confronted with the theory. Here, the data primarily serve to provide parameter values for the model. As the main conclusion of this study likely depends on these values, precise quantitative data are critical.

Specific points

1. model

The model looks fine (to the extent I can evaluate it). One assumption of the model, however, may need to be clarified. One main conclusion of this study is that Dpp recycling increases over time, directing scaling. I therefore presume that increased recycling of Dpp at the source results in increased Dpp production. Consistent with this, a previous study by Gonzalez-Gaitan and colleagues found that the amplitude of the DPP gradient increases during development. Despite this, the model appears to consider a constant rate of production from a localized source (SM, Section 1.A.1). If it is indeed so, could the authors justify this choice? And if one considers a variable rate of Dpp production, would the conclusions of the paper remain the same?

We have previously shown that the increase in amplitude is due to decrease in effective degradation rate and increase in the width of the source[14,25], whereas the measured effective production rate remains constant[14].

In the Approximate Bayesian approach to analyze the FRAP data, we fit normalized recoveries and therefore the absolute levels of the production rate do not matter and cannot affect the parameterization of the other transport rates.

Otherwise, the authors should state somewhere that the model best applies to the posterior (P) compartment of the wing disk. Indeed, in the model, cells are either part of the source (vn=1) or not (vn=0), and this parameter remains fixed. This situation corresponds to the P compartment: all P cells are vn=0, and source cells are A cells at the A-P boundary (vn=1). By contrast, the A cells close to the A-P boundary would switch from vn=1 to vn=0 as the tissue grows. This switch is not implemented in the model. This brings another question. Were all experimental data acquired in P cells? While I presume so, given the legend of Figure 2, it may be better to spell it out clearly. This is important and has several implications. For instance, one might expect that the size of the source of Dpp and of GFP-Dpp driven by Gal4/LexA might slightly differ given the stability of Gal4/LexA which could persist in the A cells that no longer express the dpp gene.

Indeed, our analysis is restricted to the target cells in the posterior compartment. This is now better clarified in the text.

> *page 2. We focused on the target cells in the posterior compartment of the wing imaginal disc, where we distinguish five pools of Dpp and consider trafficking through these pools mediated by eight different transport steps (Fig. 1d).*

2. Dpp overexpression

A limitation of this study is that Dpp transport rates are estimated using an overexpressed version of Dpp, GFP-Dpp, expressed under the control of dpp-Gal4 or dpp-LexA. Importantly, the authors made the effort to generate an endogenous version of GFP-Dpp using CRISPR/Cas9 knock-in, but this tool had two severe issues (section 2,

1.2.2): first, this dpp allele was homozygous lethal, implying that this version of Dpp is not functional. Second, one copy of GFP-Dpp was too weak to be used in the FRAP assay used to parametrize the model. This reveals that the levels of GFP-Dpp accumulation in the Gal4 (or LexA) conditions are significantly higher than endogenous Dpp levels.

We quantified the amounts of GFP-Dpp expressed by lexA in the wing discs with immunoblot and compared it to the amount of GFP-Dpp endogenously tagged and they are similar (Extended Data Figure 7a). We have now added the quantification of the immunoblot to the figure.

Whether this difference matters or not should be examined. Specifically, the authors could study the extent to which endogenous GFP-Dpp and overexpressed GFP-Dpp similarly scale (Ext Data Fig 1a-c),

We could observe scaling of the GFP-Dpp gradient using the GFP-Dpp CRISPR, expressing GFP-Dpp at endogenous levels (NEW Extended Data Figure 14a,b). The λ/L scaling factor in the CRISPR is not significantly different to the scaling factor in the GFP-Dpp^{LOP} line. We added this now to the description of the CRISPR line in the Supplementary Material.

recycle (diffusion rate of internalized and photoconverted Dendra-GBP; Fig 1a,b and Ext Data Fig 1d-f) etc…

In order to confirm that the behavior of GFP-Dpp molecules in the system endogenously expressing GFP-Dpp (GFP-Dpp^{CRISPR}) is similar to our eGFP-Dpp^{LOP}, we have performed photoconversion experiments and nanobody uptake experiments using the GFP-Dpp^{CRISPR} (NEW Extended Data Figure 14c,d). The photoconversion experiment using GFP-Dpp^{CRISPR} shows that recycled molecules contribute to the formation of the GFP-Dpp gradient (NEW Extended Data Figure 14c). Also the nanobody uptake parameters in GFP-Dpp^{CRISPR} discs are similar to nanobody uptake parameters in eGFP-Dpp^{LOP} (see NEW Extended Data Figure 14d). We have now added this to the description of the CRISPR line in the Supplementary Material.

In other words, it would be important to establish that the values of the Dpp transport parameters do not depend on Dpp levels.

Along the same line, it would be important to state somewhere that most experiments were performed in a context where GFP-Dpp competes with endogenous Dpp. And it is needed to test whether this competition affects the behavior of GFP-Dpp. One possible experiment would be to overexpress untagged Dpp and examine the effect of excess Dpp on the transport rates measured for GFP-Dpp. Since the authors propose that changes in Dpp binding rates underlie scaling, the possibility that competition and overexpression of the GFP reproter affects parameter values should be seriously considered.

We now stated in the Supplementary Material that most experiments are performed in a context where GFP-Dpp competes with endogenous Dpp.

In order to test whether the competition of unlabelled Dpp affects the behavior of GFP-Dpp, we have overexpressed untagged Dpp in a eGFP- Dpp^{LOP} expressing background, as the reviewer has suggested. We have then performed nanobody uptake and FRAP assays to examine the effect of excess Dpp on the transport rates measured for eGFP-Dpp^{LOP}. We observe no difference in the nanobody uptake and the FRAP recovery curves in the context of Dpp overexpression as compared to our experiments without overexpressing untagged Dpp (see NEW Extended Data Figure 13a-c). We have included this analysis in the new version of the Supplementary Material Section 2.

> *Note that most experiments using eGFP-Dpp^{LOP} have been done in the presence of the endogenous gene and therefore they represent an overexpression condition. To test the impact of overexpression on dynamics of Dpp transport and the formation of the gradient, we performed nanobody uptake and FRAP assays in a condition in which we overexpress untagged Dpp under the control of dppGal4 in a background of expressing eGFP-Dpp^{LOP} (Extended Data Figure 13a-c). We observe no difference in the nanobody uptake and the FRAP recovery curves in the context of Dpp overexpression as compared to our experiments without overexpressing untagged Dpp.*

3. Dpp recycling and photoconversion

This is a nice assay(Fig 1a,b). It clearly suggests that internalized Dpp can be recycled and diffuse. Since at t=0 post-conversion, a strong signal is seen at the source, recycled Dpp detected at later time points at a distance from the source likely originated from the source cells. However, whether recycling also occurs in the growing target tissue, which is critical for the spreading of Dpp, is assumed but not demonstrated here.

We have now performed new experiments, photoconverting GBP-Dendra2 in the target tissue, as the reviewer has suggested (NEW Figure 1b).In these experiments we observe that photoconverted endocytosed GBP-Dendra2* molecules can spread from the target tissue to form a gradient in the anterior compartment. This experiment

demonstrates, that recycling and unbinding occurs also in the target tissue. We have included this experiment in the new version of the main text of the manuscript

> *page 2. Photoconversion of this intracellular GFP-Dpp-bound GBP-Dendra2 in a region of interest (ROI; whether in the source (Fig. 1a) or target tissue (Fig. 1b)) allows us to address whether these internalized molecules can move to neighbor cells and form a gradient there (for details and controls of this assays, see SM 2.8; Extended Data Figures 5e-g, 8a-f).*

and

> *page 2. Because photoconversion in the target territory also lead to formation of a gradient (Fig. 1b), this shows that Dpp molecules can be re-exocytosed also from the target cells.*

More importantly, it is not clear whether recycling is actually important for the formation of the Dpp gradient. Is it possible to experimentally up(down)-regulate the recycling of Dpp and examine the consequences on the Dpp gradient? (I must say that I have no good suggestion to offer here)

In order to downregulate recycling of Dpp, we have downregulated the expression of the recycling Rab proteins: Rab11 and Rab4. We performed RNAi by expressing their corresponding dsRNA in the posterior compartment of the wing disc in a temperature dependent manner for 4 days before dissection. We have observed a decrease of the decay length of eGFP-Dpp^{LOP} gradient (see NEW Extended Data Figure 12). A decrease of the Dpp decay length is expected in this condition, as the recycling plays an important role in gradient formation. We have added the description of this experiment to the Supplementary Material Section 2.

> *We show that large discs are in a combined transport regime, where extracellular diffusivity, recycling and unbinding are all important for the formation of the gradient. Consistently, the boosts B_{k_r}, $B_{k_{off}}$ and B_{D_0} are important (Extended Data Figure 3c). Consistent with a large B_{k_r}, the decay length of the gradient is shortened upon partial silencing of recycling by RNAi through expression of dsRNA for the recycling Rabs (Rab4 and Rab11) in the posterior target cells (Extended Data Figure 12).*

Otherwise, I do not understand how averaging intensity values in three independent experiments/discs (n=3) can produced high amplitude variations along the x-axis (see t=87.5min data plots). Please comment.

The intracellular fraction in endosomes dominates the profile (87% of the signal is endosomal, Figure 2g). Since these are discrete objects, they appear as spatial fluctuations ("amplitude variations along the x-axis") in the profile. These variations are not exclusive of this experiment, but can also be seen in normal GFP-Dpp gradient profiles when a reduced sample size is considered. Indeed, in the previous version of the manuscript, we showed the average of only three gradients. In the revised version, we have increased our sample size and therefore the gradients look smoother and more similar to other GFP-Dpp gradient profiles shown in this and other reports (NEW Figure 1c).

4. scaling and data selection

The data shown in Fig 2a are both nice and important. In this plot, the decay length values are given for different binning of length values. This is unusual.

In the previous version of the plot in Figure 2a, the binning was different for control and *pentagone* mutant conditions. In the revised version, the binning is now the same for both.

Would it be possible to plot instead the raw data as a cloud of individual data points?

Raw data is presented in Extended Data Figure 6a. We now added a reference to this figure in the legends of Figure 2a.

I am asking this because I did not understand why the authors chose to bin the decay values in two categories corresponding to 144.6 +/-4 and 79.3 +/-1.4

The choice of 144 µm and 79 µm is motivated by the fact that the smallest discs in which the assays can reliably be performed are about 80 µm and in discs larger than 144 µm, the amplitude of the Dpp gradients start normally a rapid decrease in preparation for pupariation. In the new version of the Supplementary Material we have clarified this point.

In addition, there might be a confusion about the structure of our sample. We choose discs whose posterior lengths are close to these two sizes (144 µm and 79 µm, with variabilities smaller than 10% those sizes). We did not consider a data set where discs are continuously changing in size from 144 µm to 79 µm and bin it in two categories.

(see legend of Fig 2a; note that different SEM values are given in Ext Data Fig 1b and Table 2; please clarify).

This was a typo: the values given in Table 2 were standard deviations (different to SEM in the figure legend). This has now been corrected.

The motivation for this very specific binning remains unexplained. Specifically, I am concerned that the corresponding decay length values (28.1 +/-1 and 14.9 +/-0.5; Table S2) may not fit well on a linear regression of the data plotted in Fig 1a, suggestive of data selection.

First, the regression was performed on the raw data before binning (Extended Data Figure 6a). The characteristic decay length for discs of 80 and 144 µm was estimated from this linear fit. Binning was not utilized for the linear regression and therefore "data selection" is not an issue here. Binning was only performed for ease of representation.

For this reason, I would strongly recommend avoiding categorization whenever possible. I do understand, however, that categorization cannot be avoided when performing transport assays. However, the size distribution of the experimental discs needs to be reported. Currently, this information is not easy to find (or not provided). For instance, in Fig 2e, the authors used l=144 simply to name a sample category. This naming cannot replace the mean/sem values corresponding to this particular sample. Please provide this information (in this and all other experiments).

The posterior compartment length (ℓ) with the standard errors of the mean are given in the legend of Figure 2 as well as in the bar plot of the Extended Data Figure 1b. We have now added the exact average posterior compartment length for large discs together with the SEM to the main text of the manuscript.

5. nanobody internalization assay

This nice assay measures the rate of endocytosis of GFP-Dpp bound to a nanobody,

This assay does not only estimate the rate of endocytosis. The dynamics of nanobody uptake depend on the rate of endocytosis, recycling rate and k_o rate (a sum of the immobilization and the degradation rates from the early endosome). In Supplementary Material Section 2 chapter 2.2.1 we detail the complete theory of how the dynamics of the uptake depends on those rates. In particular the uptake has two superimposed kinetics: exponential and linear. The exponential kinetics (p) depends on k_o and k_r (equation S2.18). the dynamics of the linear regime (A) depends on the endocytosis rate (k_N), the recycling rate (k_r) and k_o.

The linear part is due to the fact that Alexa555 (a synthetic, not natural molecule) is not degraded in endosomes in the time scale of the experiment. Therefore instead of reaching a plateau, the signal of uptaken molecules increase linearly (Extended Data Figure 1f).

following a binding step at 4°C (note: please state in section 2, 2.2.2, that discs are incubated with the same concentration of GBP-Alexa555 during imaging at 25°C, hence allowing for fluid phase endocytosis; if GBP-Alexa555 is not added, fluid-phase endocytosis may be underestimated).

See below for a discussion of the fluid phase issue.

While this assay aims at quantifying endocytosis at steady-state, it actually measures a pulse of Dpp internalization following transfer of the disc at 25°C.

This is not a pulse, as we do not wash away the GBP-Alexa555 after the 4°C binding step. Instead of a pulse, there is a continued uptake of GBP-Alexa555 bound to GFP-Dpp during the 25°C step.

This is an issue. Indeed, after 30 min of incubation, rapid Dpp internalization is no longer detected and only fluid-phase endocytosis is monitored for another 30 min (the linear regime of internalization corresponds to fluid-phase endocytosis, as shown in Ext Data Fig 7i). In sum, this assay may not monitor Dpp endocytosis at steady state. Surprisingly, a 5-fold variation in fluid-phase endocytosis (phenomenological parameter A in Ext Data Fig 1f) was observed between early and late discs (Ext data Table 2: 0.0011 vs 0.00021). This 5-fold variation likely reflects a combination of technical and biological noises.

In the GFP-Dpp expressing discs, the linear regime in the nanobody uptake experiment (captured by the slope A) does not correspond to the fluid phase uptake, but depends on k_N, k_r and k_o: $A = \frac{k_N k_o}{k_o + k_r}$ (see Supplementary Material Section 2 chapter 2.2.1).

Fluid phase internalization of the nanobody is defined by the nanobody molecules internalized passively, hijacking existing endocytosis events in the absence of GFP-Dpp. When we incubate a disc without GFP-Dpp expression with GBP-Alexa555 at the concentration used in the uptake experiments with GFP-Dpp, fluid phase uptake is

negligible with respect to nanobody internalization in the presence of GFP-Dpp (see Extended Data Figure 7g,h). The fact that this fluid phase is negligible indicates that internalization of GFP-Dpp bound to the nanobody is receptor-mediated.

The confusion might have come from the Extended Data Figure 7j, which shows a significant fluid phase slope. This experiment was done with 10 fold the concentration used in the uptake experiments on GFP-Dpp discs to illustrate robustly the exclusively linear (not exponential) regime of the fluid phase uptake, where recycling does not take place.

On a related note, the accumulation profile of internalized nanobody appeared somewhat variable (compare Figs 2c, 4b and 4d): is variability due to biological or technical variations?

The difference in the accumulation profile of the internalized nanobody, that the reviewer notices, results from a difference in the duration of the nanobody incubation. Indeed, in the previous version, we already specified that in Figure 2c nanobody uptake took place for 166 minutes (see figure) while in Figure 4d uptake took place for 30 minutes (see legend). This was because in the iFRAP experiment (Figure 4d), we wanted to avoid the filling of the immobile pool, which cannot contribute to recycling. This was explained in Supplementary Material Section 2 chapter 2.7.1.

Otherwise, what are the samples (n=15) shown in Ext Data Fig 1d?

The nanobody uptake assay was performed in 13 discs. Data points and the fit curves for these 13 discs is color coded. The fit to the uptake dynamics was not performed one by one in each of the 13 discs, but simultaneously for all the 13 samples to achieve a more robust determination of the parameters A, B and p. This is explained in the Supplementary Material Section 2 chapter 2.2.3.

6. iFRAP

FRAP was performed on the internalized pool of GBP-Alexa555 (Fig 4c-g). Internalization, following a 4°C binding step, was performed for 35 min, ensuring that a large fraction of internalized GBP-Alexa555 entered in GFP-Dpp dependent manner. Following photobleaching, fluorescence recovery was monitored over the next 30 min. Importantly, at this stage (35-60min of incubation), GBP-Alexa555 was reported to be mostly endocytosed in a fluid phase manner.

As explained above, the late phase of recovery in the assay (after 35 minutes) does not correspond to internalization of GBP-Alexa555 through fluid phase endocytosis. And fluid phase internalization is negligible at the nanobody concentrations used in this assay.

Instead, the iFRAP experiment shows the dynamics of GBP-Alexa555 molecules bound to GFP-Dpp, belonging exclusively to the pool of internalized molecules (because of the acid wash performed prior to photobleaching) which are then recycled back, detach from the plasma membrane, diffuse and appear in neighboring cells in the bleached ROI.

Thus, whether this assay monitors the receptor-dependent endocytosis of a GFP-Dpp/GBP-Alexa555 complex remains to be established. It is therefore unclear whether this assay actually monitors the recycling, diffusion and endocytosis of GFP-Dpp.

It is worth noting again that fluid phase internalization is negligible. Indeed we only detect internalization of GBP-Alexa555 bound to GFP-Dpp as we explained above, which we know because of the controls we performed to exclude fluid phase endocytosis (Extended Data Figures 7g,h); fluid phase is defined here as GBP-Alexa555 internalization which is independent of GFP-Dpp and therefore also independent of the Dpp receptor.

These facts unambiguously exclude fluid phase internalization as an issue. A clear indication of this is that upon incubation with GBP-Alexa555, the spatial pattern of internalized GBP-Alexa555 in the disc corresponds to a gradient with the same decay length as GFP-Dpp (Extended Data Figure 7m,n). Fluid phase would give a flat profile which would not parallel the GFP-Dpp gradient.

7. Transcytosis

Previous work by Gonzalez-Gaitan has suggested that transcytosis is required for the formation of the Dpp gradient (Entchev et al., Cell, 2000). Here, Gonzalez-Gaitan suggests that transcytosis plays no significant role (Fig 3b). Given that the original finding was very controversial, this new conclusion by the same authors should more clearly appear in the main text. Currently, a discussion on the possible role of transcytosis is buried in section 2, 4.1.3, and will only be accessible to a very limited audience.

In the paper mentioned by the reviewer (Entchev *et al.*[26]), the authors find that the range of Dpp signaling is controlled by endocytic trafficking and degradation. They proposed three scenarios that are consistent with the experimental data of the paper. Transcytosis is only one of them (text from the Entchev paper):

> *The Rab mutant analysis reveals that the range of Dpp signaling is controlled by endocytic trafficking and degradation in the receiving cells. This finding is consistent with three possible scenarios which do not exclude each other. (1) The Dpp receptor is downregulated by degradation in the DRab7 gain-of-function mutant, whereas its recycling is impaired or enhanced in the DRab5 dominant-negative and overexpression conditions, respectively (…) (2) Dpp internalization by endocytosis carries the ligand to an endosomal compartment where the actual signal transduction takes place. (…) (3) Dpp travels intracellularly through the endocytic pathway through planar transcytosis.*

At that point, the authors did not favor any of these possibilities in particular. However, the data then made us commit to a scenario which our manuscript now seems to support (text from the Entchev paper):

> *We favor a model where the balance between recycling and degradation of the ligand in the endocytic pathway determines the shape of the gradient*

Thus, in the Entchev, 2000 paper, the key phenomena which are emphasized to be important for Dpp gradient formation are recycling and degradation. Our current work is in agreement with this finding: in large discs (and to a lesser extent also in small discs), Dpp molecules that are endocytosed and recycled back to the plasma membrane are crucial contributors to the diffusing extracellular pool that forms the gradient (the role of the recycling module).

8. DppTimer

The green/red ratio is interpreted to reflect the age of Dpp. However, this ratio also depends on intracellular pH. Following internalization into acidic endosomal compartments, the fluorescence signal of sfGFP (but not those of mKate2) is quenched. Thus, this approach cannot be used to directly infer the ages of the Dpp molecules.

The concern raised by the reviewer regarding the quantification of the age ratio using the Dpp^{Timer} is important. In this analysis, a precise estimation of the ratio of sfGFP and mKate2 fluorescence is essential (see Supplementary Material Section 2 chapter 2.6 for the description of the assay). Indeed, a large portion of Dpp is in endosomal compartments inside cells, where pH is lower than 7. Acidic pH could in principle affect to a different extent the fluorescence of the two different fluorophores of the Timer molecule: sfGFP and mKate2.

In order to address this issue, we have performed the following experiments:

1) *In vivo,* we estimated the effect of pH on fluorescence of the sfGFP and mKate2 in the Dpp^{Timer} molecule in wing imaginal discs. We did this by neutralizing the pH to 7 in the endosomal compartments by means of bafilomycin treatment (NEW Extended Data Figure 9h,i) and compared the fluorescence intensity of the two fluorophores before and after neutralizing the pH. To confirm the efficiency of bafilomycin in our treatment, we followed *in vivo* the pH of endosomes with LysoSensorTM (ThermoFischer Scientific): before bafylomicin treatment, discs were incubated in a solution of 1µM of LysoSensorTM in Clone 8 medium for 30 minutes. Acidotropic LysoSensorTM accumulated in acidic intracellular compartments (NEW Extended Data Figure 9g). Following neutralization of pH upon bafylomicin treatment, LysoSensorTM fluorescence was decreased and no more accumulation in intracellular compartments can be observed, NEW Extended Data 9g).
2) *In vitro,* we estimated the effect of pH on sfGFP and mKate2 fluorescence intensity using a purified Timer molecule (NEW Extended Data Figure 9j):

In vivo experiments (#1) show that the effect of pH in Dpp-containing endosomes is similar on both fluorophores: at the acidic pH, fluorescence intensity decays by $62 \pm 4\%$ and $59 \pm 7\ \%$ for sfGFP and mKate2, respectively (NEW Extended Data Figure 9h). Because both fluorophores are affected in parallel and the age measurement with the Timer is based on the ratio of their fluorescence, this result indicates that the acidic pH of the endosomes does not compromise the validity of the Timer.

Indeed according to the literature[27,28], the pKa of sfGFP and mKate2 are very similar, 5.9 and 5.4, respectively. At pH=pKa half the molecules are protonated and it is known that protonation of these fluorophores quenches the fluorescence *i.e.* at pH=pKa half of the fluorescence is expected to be lost. The observed decrease in fluorescence upon neutralization *in vivo* corresponds therefore to an average pH of 5.3-6 in Dpp-containing intracellular compartments, which is in the published range of pH from early (6.2) to late (5.0) endosomes[29]. Consistently, our

in vitro experiments (#2) confirm that, within this interval (5.0-7), the change of fluorescence intensity of sfGFP and mKate2 with pH is also similar.

9. Pentagone function

The data shown in Ext Data Fig 1c suggests that the Dpp gradient cannot be approximated to a simple exponential. Is this observation reproducible? Could the authors quantitatively examine how well the experimental data fit with an exponential in both wild-type and pentagone mutant discs? In other words, does the loss of Pentagone affect the shape of the Dpp gradient? This needs to be examined.

We have generated a new dataset of eGFP-Dpp^{LOP} gradients in *pent*^{2} l=130 µm discs. We have fitted the gradients to an exponential function and show now a plot of the residuals of the fit as compared to the corresponding residuals in control l=144 µm discs. We do not observe any significant differences in the quality of the exponential fit between the control and the mutant condition (NEW Extended Data Figure 1d,e).

References

1 Kicheva, A. *et al.* Kinetics of morphogen gradient formation. *Science* **315**, 521-525, doi:10.1126/science.1135774 (2007).

2 Belenkaya, T. Y. *et al.* The retromer complex influences Wnt secretion by recycling wntless from endosomes to the trans-Golgi network. *Dev Cell* **14**, 120-131, doi:10.1016/j.devcel.2007.12.003 (2008).

3 Lucken-Ardjomande Hasler, S., Vallis, Y., Pasche, M. & McMahon, H. T. GRAF2, WDR44, and MICAL1 mediate Rab8/10/11-dependent export of E-cadherin, MMP14, and CFTR DeltaF508. *J Cell Biol* **219**, doi:10.1083/jcb.201811014 (2020).

4 Bowen, A. B., Bourke, A. M., Hiester, B. G., Hanus, C. & Kennedy, M. J. Golgi-independent secretory trafficking through recycling endosomes in neuronal dendrites and spines. *Elife* **6**, doi:10.7554/eLife.27362 (2017).

5 Stapornwongkul, K. S., de Gennes, M., Cocconi, L., Salbreux, G. & Vincent, J. P. Patterning and growth control in vivo by an engineered GFP gradient. *Science* **370**, 321-327, doi:10.1126/science.abb8205 (2020).

6 Akiyama, T. *et al.* Dally regulates Dpp morphogen gradient formation by stabilizing Dpp on the cell surface. *Dev Biol* **313**, 408-419, doi:10.1016/j.ydbio.2007.10.035 (2008).

7 Fujise, M. *et al.* Dally regulates Dpp morphogen gradient formation in the Drosophila wing. *Development* **130**, 1515-1522, doi:10.1242/dev.00379 (2003).

8 Guha, A., Sriram, V., Krishnan, K. S. & Mayor, S. Shibire mutations reveal distinct dynamin-independent and -dependent endocytic pathways in primary cultures of Drosophila hemocytes. *J Cell Sci* **116**, 3373-3386, doi:10.1242/jcs.00637 (2003).

9 Gupta, G. D. *et al.* Analysis of endocytic pathways in Drosophila cells reveals a conserved role for GBF1 in internalization via GEECs. *PLoS One* **4**, e6768, doi:10.1371/journal.pone.0006768 (2009).

10 Mayor, S., Parton, R. G. & Donaldson, J. G. Clathrin-independent pathways of endocytosis. *Cold Spring Harb Perspect Biol* **6**, doi:10.1101/cshperspect.a016758 (2014).

11 Sabharanjak, S., Sharma, P., Parton, R. G. & Mayor, S. GPI-anchored proteins are delivered to recycling endosomes via a distinct cdc42-regulated, clathrin-independent pinocytic pathway. *Dev Cell* **2**, 411-423, doi:10.1016/s1534-5807(02)00145-4 (2002).
12 Hemalatha, A., Prabhakara, C. & Mayor, S. Endocytosis of Wingless via a dynamin-independent pathway is necessary for signaling in Drosophila wing discs. *Proc Natl Acad Sci U S A* **113**, E6993-E7002, doi:10.1073/pnas.1610565113 (2016).
13 Norman, M., Vuilleumier, R., Springhorn, A., Gawlik, J. & Pyrowolakis, G. Pentagone internalises glypicans to fine-tune multiple signalling pathways. *Elife* **5**, doi:10.7554/eLife.13301 (2016).
14 Wartlick, O. *et al.* Dynamics of Dpp signaling and proliferation control. *Science* **331**, 1154-1159, doi:10.1126/science.1200037 (2011).
15 Aguilar Hidalgo, D., Hadjivasilou, Z., Romanova-Michaelides, M., González-Gaitán, M. & Jülicher, F. Dynamic modes of morphogen transport. arXiv:1909.13280 (2019). <https://ui.adsabs.harvard.edu/abs/2019arXiv190913280A>.
16 Zhou, S. *et al.* Free extracellular diffusion creates the Dpp morphogen gradient of the Drosophila wing disc. *Curr Biol* **22**, 668-675, doi:10.1016/j.cub.2012.02.065 (2012).
17 Fried, P. & Iber, D. Dynamic scaling of morphogen gradients on growing domains. *Nat Commun* **5**, 5077, doi:10.1038/ncomms6077 (2014).
18 Zhu, Y., Qiu, Y., Chen, W., Nie, Q. & Lander, A. D. Scaling a Dpp Morphogen Gradient through Feedback Control of Receptors and Co-receptors. *Dev Cell* **53**, 724-739 e714, doi:10.1016/j.devcel.2020.05.029 (2020).
19 Ben-Zvi, D., Pyrowolakis, G., Barkai, N. & Shilo, B. Z. Expansion-repression mechanism for scaling the Dpp activation gradient in Drosophila wing imaginal discs. *Curr Biol* **21**, 1391-1396, doi:10.1016/j.cub.2011.07.015 (2011).
20 Hamaratoglu, F., de Lachapelle, A. M., Pyrowolakis, G., Bergmann, S. & Affolter, M. Dpp signaling activity requires Pentagone to scale with tissue size in the growing Drosophila wing imaginal disc. *PLoS Biol* **9**, e1001182, doi:10.1371/journal.pbio.1001182 (2011).
21 Belenkaya, T. Y. *et al.* Drosophila Dpp morphogen movement is independent of dynamin-mediated endocytosis but regulated by the glypican members of heparan sulfate proteoglycans. *Cell* **119**, 231-244, doi:10.1016/j.cell.2004.09.031 (2004).
22 Vuilleumier, R. *et al.* Control of Dpp morphogen signalling by a secreted feedback regulator. *Nat Cell Biol* **12**, 611-617, doi:10.1038/ncb2064 (2010).
23 Beckett, K. *et al.* Drosophila S2 cells secrete wingless on exosome-like vesicles but the wingless gradient forms independently of exosomes. *Traffic* **14**, 82-96, doi:10.1111/tra.12016 (2013).
24 de Madrid, B. H., Greenberg, L. & Hatini, V. RhoGAP68F controls transport of adhesion proteins in Rab4 endosomes to modulate epithelial morphogenesis of Drosophila leg discs. *Dev Biol* **399**, 283-295, doi:10.1016/j.ydbio.2015.01.004 (2015).
25 Aguilar-Hidalgo, D. *et al.* Critical Point in Self-Organized Tissue Growth. *Phys Rev Lett* **120**, 198102, doi:10.1103/PhysRevLett.120.198102 (2018).
26 Entchev, E. V., Schwabedissen, A. & Gonzalez-Gaitan, M. Gradient formation of the TGF-beta homolog Dpp. *Cell* **103**, 981-991, doi:10.1016/s0092-8674(00)00200-2 (2000).
27 Roberts, T. M. *et al.* Identification and Characterisation of a pH-stable GFP. *Sci Rep* **6**, 28166, doi:10.1038/srep28166 (2016).

28 Shcherbo, D. *et al.* Far-red fluorescent tags for protein imaging in living tissues. *Biochem J* **418**, 567-574, doi:10.1042/BJ20081949 (2009).
29 Hu, Y. B., Dammer, E. B., Ren, R. J. & Wang, G. The endosomal-lysosomal system: from acidification and cargo sorting to neurodegeneration. *Transl Neurodegener* **4**, 18, doi:10.1186/s40035-015-0041-1 (2015).

Reviewer Reports on the First Revision:

Referee #1 (Remarks to the Author):

Romanova-Michaelides et al have extensively revised their study and clarified the points raised in the initial review.

The strength of the study is the rigorous measurement of biophysical mechanisms of ligand spread in the well characterized Drosophila wing disc. The additional data and rebuttal answer most of the questions I raised during the initial review.

The study will be of interest to the field and the approaches the authors use will be broadly appreciated. I have three remaining points:
- I am not convinced that the short format of the paper is the most appropriate for this type of study. Almost 90% of the material ends up in Supplemental appendices and the main body of the paper is little more than an extended abstract for the actual details of the study. I accept the authors are constrained by the journal format, but it does make the study appear less scholarly than it is.

- I take on board the authors' arguments about loss of Dpp ligand from the epithelium. However, this does appear to be one of the weaker aspects of the study. The authors' place a lot of weight on the behavior of secreted GFP in the recently published Stapornwongkul et al study. Whether Dpp has the same characteristics as extracellular GFP is an assumption. Moreover, the experimental test the authors perform of whether or not there is loss of Dpp molecules to the extracellular space is carried out at 4oC in fairly artificial conditions.

- Lack of knowledge of the molecular identity of the putative internalization receptor implicated by the analysis is a potential weak spot. If this receptor was identified, manipulation of its expression would provide a convincing test of the authors' model. Moreover, the apparent absence of evidence of what this receptor might be is troublesome. Given the intensity with which this system has been studied, candidates from genetic screens might be expected because mutations would be predicted to change the Dpp gradient and therefore wing patterning.

Referee #3 (Remarks to the Author):

I appreciate the thorough and well-thought responses of authors to my concerns. I am fully satisfied with their arguments and with the additional data provided to support these. A small but very important issue, though. I wonder why the transcytosis module in their original equation. There is no convincing experimental evidence supporting this mode of Dpp movement in this particular model system. Could authors briefly argue why? Is this module absolutely necessary? Do their data fit removing it? If so, why not doing so?

Referee #4 (Remarks to the Author):

The authors have revised their manuscript by providing more data as well as clarifications.

Revisions, however, did not significantly change some critical aspects of this study (see points 1-4 below). I also agree with referee #1 when she/he writes that "it is difficult to assess the rigorousness of the experiments and the confidence one should have in the conclusions" (referee #1). This somewhat echoes some of the points that I raised.

Otherwise, a recent study of an engineered GFP gradient from the lab of JP Vincent, cited as ref #4 in the manuscript, showed that diffusible GFP can replace Dpp in the presence of Dpp receptors rewired to bind GFP. This suggests that regulated trafficking of the ligand may not be critical to form a scaling gradient. The authors further showed that "concomitant expression of glycosylphosphatidylinositol (GPI)–anchored non-signaling receptors further improved patterning, to near wild-type quality. Theoretical arguments suggest that GPI anchorage could be important for these receptors to expand the gradient length scale while at the same time reducing leakage" (abstract). This paper is fully relevant for this study since it provides experimental evidence indicating that tissue-level diffusion of Dpp and GPI-anchored non-signaling binders may be sufficient to extend the gradient with no need for specific regulation of Dpp trafficking. Despite its relevance, this study is first cited to indicate that "unbound ligand molecules could in principle leak out of the epithelium". This is not appropriate.

I first comment below the revisions addressing the points raised in my review:

1. I maintain that the main claims of the paper (scaling is achieved by increased recycling of Dpp, driven by Pentagone-dependent changes in the binding rates of Dpp to its internalization receptor, see abstract) are only supported by indirect evidence. As indicated earlier, the internalization receptor of Dpp is unknown and binding rates can only be indirectly inferred.

2. The authors have taken into consideration my comment about the photoconversion of internalized Dpp and now acknowledge that this assay addresses whether internalized Dpp CAN contribute to gradient formation. In this revised version, the authors provide evidence that Dpp CAN be re-exocytosed from the target tissue (Fig 1b, new panel). Of note, i) no quantification is provided and ii) formation of the gradient is observed over the source. Hpowever, the extent to which internalized Dpp DOES contribute to gradient formation remains unaddressed. Despite this, the authors did not change their conclusion ("therefore recycling contributes to gradient formation"). Thus, the issue remains.

3. "The experimental data are not meant to challenge the model but to make it work". The point I raised about the model and its parametrization echoes the first comment of referee #2 ("observations and conclusions are a consequence of the particular modeling assumptions"). The authors respond to our criticisms by stating that they "performed here a minimal description of phenomena that we know happen in the tissue" (p. 18) and that their "approach is not dependent on a particular model, but instead consists of developing a phenomenological description that takes into account processes that we know occur in the tissue" (p. 6). They further state that "the transport of Dpp with 5 pools and 8 rates is not a model, it is a physical description" and that "these effective pools and rates do exist". These statements are somewhat puzzling and overall naive. This study provides a good example of forward modeling based upon known, or suspected, causalities ("processes that we know occur in the tissue: diffusion, binding etc…") that are expressed in the form of a set of equations. As with any other model, the model proposed here is based on a set of assumptions. Some of these assumptions are clearly spelled out (gradient at steady-state, 5 pools of Dpp with various transport mechanism), others are hidden in the model (e.g. no leakage of Dpp into the extracellular space). Since the model studied here is derived from phenomenology, the model is essentially a test of its assumptions and must be falsifiable (on this topic, I encourage the authors to read the review by J Gunawardena https://doi.org/10.1186/1741-7007-12-29). In this study, the authors have not attempted to falsify their model but have parametrized a model to fit their experimental data.

4. Precision of the measurements

While the authors agree that precise measurements are important for the "robustness of their final conclusion", they actually refer to SM sections 2.5.2 (that discusses how the confidence intervals were established for k_N, k_r and k_o) and 2.3.7 (that examines how relaxing threshold values for the calculated extracellular pool and for the fitting of the FRAP data) to indicate that their conclusions are robust to changes in the corresponding measured values. This may be seen as paradoxical. I still find difficult to evaluate how precise and reproducible are the results of the transport assays developed in this study.

5. Pentagone levels over time
The authors used a pent-lacZ transgene to examine the expression of Pentagone. This approach is sub-optimal as it only reports upon the activity of a 2 kb fragment of regulatory DNA, whereas FISH is an easy and quantitative approach to detect endogenous gene expression. Otherwise, it is not clear how discs area (Extended Data Fig 15b) relates to the l=80 and l=144 stages examined in the rest of the paper. Based on the data plotted in Ext Data Fig 11e, it appears that not a single disc in Extended Data Fig 15b corresponds to l=144. Overall, these data are inconclusive.

6. model
I must admit that I do not understand the authors response: does recycling increase at the source? If so, how does production rate remain constant despite an increase in recycling? If not, what kind of mechanism might prevent an increase in Dpp recycling only at the source?

7. Dpp overexpression
The authors have addressed my concern and now show that GFP-Dpp produced at the locus (CRIPSR) scale and recycle similarly as GFP-Dpp (LOP) (Ext Fig 14). Also, the authors now suggest that over-expression of Dpp has no significant effect on the behavior of GFP-Dpp (LOP). Note, however, that the relative levels of Dpp associated with these two genotypes were not determined.

8. Dpp recycling and photoconversion
See point #2 above.

9. scaling and data selection
- Fig 2a was revised
- raw data are now plotted in Ext Fig 6a, showing the linear regression

10. nanobody internalization assay/iFRAP/transcytosis/Dpp timer/Shape of the gradient in Pentagone mutants
The authors have clarified the issues raised in my review

I also provide below a few other comments about issues raised by referees #1-3

referee #1
- p. 3, Fig 1a,b: the authors did not address the point raised by the referee and maintains their conclusion: "photoconverted molecules were ONLY in the endocytic pathway" (also Refree #2, p.11).

- p.3-4, leakage into the extracellular space: the absence of leakage is one assumption of the model. The authors argue that the formation of the Dpp gradient is unlikely to be influenced by the loss of molecules in the extracellular space based on a recent paper by J.-P. Vincent and colleagues (2020) who actually showed the opposite for extracellular GFP: "in conclusion, a single extracellular binding species reveals the gradient of an inert protein in vivo, but leakage in the hemolymph occurs and interferes with the gradient's shape, most obviously far from the source, at the tail end of the gradient." The critical point made by the authors in their rebuttal is that leakage is an issue only for the ligands that are secreted baso-laterally. Since signaling Dpp is apical, leakage is not considered to be a significant issue: "GFP-Dpp is present in the apical-most 5mic of

the epithelium, where there is no leakage, we did not consider leakage as a major phenomenon in the system." While I understand the logic of this argument, whether GFP-Dpp leaks into the extracellular space remains an assumption of the model that is not sufficiently well tested. In Ext Data Fig 13, the fraction of GBP-Alexa555 leaking into the extracellular fraction is not measured and no positive control is included.

referee #2
I agree with the referee when she/he writes that "most rates can be measured only indirectly" and that "inferred values are model-dependent". This echoes my own evaluation ("the rate of Dpp recycling was not measured", "the rates of Dpp binding to its receptors were not directly measured", "analysis of the binding rates of Dpp to its putative internalization receptor cannot be performed", "because most of these transport values (Ext table 1) are not experimentally accessible" etc…). As indicated above, the idea that this study provides a phenomenological description that is not dependent on a particular model seems to be naive. I am not sure to understand why the authors insist that their conclusion is not the consequence of modeling assumptions.
One key assumption discussed by the referee is that the gradient is at steady-state. The authors now provide FRAP data to indicate that fluorescence recovery (by Dpp transport) is much faster than growth. Specifically, in Ext Data Fig 11i, the decay length of the total pool measured prior to bleaching appeared to be very similar to the decay length measured after 30 min of recovery (variations in decay length presumably reflect variations in disc size). Thus, the assumption of a steady-state gradient is a reasonable one. Of note, the authors state (p. 4): "this justifies the quasi-state assumption for the decay length underlying the parametrization". Again, it would probably be more correct to conclude on an assumption of the model.

Author Rebuttals to First Revision:

Referee #1 (Remarks to the Author):

Romanova-Michaelides et al have extensively revised their study and clarified the points raised in the initial review.

The strength of the study is the rigorous measurement of biophysical mechanisms of ligand spread in the well characterized Drosophila wing disc. The additional data and rebuttal answer most of the questions I raised during the initial review.

The study will be of interest to the field and the approaches the authors use will be broadly appreciated. I have three remaining points:

We thank the reviewer for his/her careful evaluation of the revised version of our MS. We are pleased that we were able to clarify all points raised in the initial review and that our manuscript is judged of interest to the field and of broad appreciation. We are confident that we can also address the three remaining points.

- I am not convinced that the short format of the paper is the most appropriate for this type of study. Almost 90% of the material ends up in Supplemental appendices and the main body of the paper is little more than an extended abstract for the actual details of the study. I accept the authors are constrained by the journal format, but it does make the study appear less scholarly than it is.

We would be happy to present our work in a longer format that can make the scholar depth of the analysis more apparent, if encouraged to do so by the editor.

- I take on board the authors' arguments about loss of Dpp ligand from the epithelium. However, this does appear to be one of the weaker aspects of the study. The authors' place a lot of weight on the behavior of secreted GFP in the recently published Stapornwongkul et al study. Whether Dpp has the same characteristics as extracellular GFP is an assumption. Moreover, the experimental test the authors perform of whether or not there is loss of Dpp molecules to the extracellular space is carried out at 4oC in fairly artificial conditions.

We agree with the reviewer that the paper by the group of Jean Paul Vincent (Stapornwongkul et al.) (JPV paper) presents an artificial situation and it is unclear whether the leakage behaviour of GFP recapitulates the behaviour of Dpp. We are therefore being very cautious in applying the conclusions of the paper to GFP transport and GFP gradient formation to Dpp. Indeed secreted GFP proteins, are unlikely to behave like Dpp itself, but could be used to infer simple properties of diffusion and leakage. However, the particular features of secretion of this secreted GFP, including the place where molecules are secreted, can indeed be confounding factors in this case. In the case of sGFP in the JPV paper, because their approach is synthetic biology, sGFP was perhaps purposely constructed such that it does not have anything to do with Dpp. Because we are interested in Dpp itself, we have now generated a new sGFP which shares many of the secretion properties of Dpp (unlike the sGFP in the JPV paper; see below). We used this construct to get an upper bound for the leakage rate of a molecule which shares secretion properties with Dpp. We use this upper bound to study the impact of considering this level of leakage on our parameterization of the transport rates.

To be more concrete, to address the contribution of leakage, in the new version of the manuscript, we show results of additional experiments which were performed at physiological temperatures as requested by this reviewer, instead of performing them at 4˚C as in the previous version. This is done with a new sGFP reagent which we have generated for the revised manuscript. We also introduced a new theoretical study which incorporates leakage to address points raised by another reviewer. In summary, we did two things: i) we have generated a new reagent ($sGFP^{Dpp}$), where secreted GFP is transcribed, translated, cleaved in the Golgi apparatus by Furin and secreted like Dpp, unlike the Patched/Wingless-based secreted GFP in the JPV paper. This new $sGFP^{Dpp}$ shows clearly that, when secreted like Dpp in the apical side, this molecule does not leak and ii) we have performed theoretical analysis where leakage is included in our 5 dynamic equations. This analysis shows that, with the leakage rate estimated for $sGFP^{Dpp}$, our conclusions are not affected: the changes in the values of the different transport rates caused by leakage are negligible and the importance of the recycling module remains unaffected.

This data strengthens the conclusion in the previous version of our manuscript that leakage is not important for Dpp gradient formation. We discuss below these two points in detail:

i) It is well established, that in the wing epithelial cells, different membrane proteins (HSPGs *vs.* Notch *vs.* EGF, etc) traffic in the secretory pathway through different, non-overlapping populations of ER/Golgi organelles[1]. The targeting of these proteins to apical versus basal secretory compartments affects the location (apical vs basolateral) in which those molecules are secreted. The specific targeting of mRNA to a subset of ER/Golgi compartments is encoded in non-translated 5' and 3' UTR sequences in the RNA as well as on signal peptide sequences at the protein level[2].

The secreted GFP reagent in the JPV paper was generated by using the 5'UTR of *patched* fused to the coding region of the signal peptide of Wingless followed by the coding sequence of GFP downstream (we term this here sGFP$^{Ptc\text{-}Wg}$; see scheme in NEW Extended Data Figure 13d). This construct is inserted in the endogenous *patched* gene and is therefore transcribed like *patched*. The authors barely see sGFP$^{Ptc\text{-}Wg}$ in the epithelium (see Figure 1B,C in the JPV paper) from what they conclude that these molecules are secreted basolaterally and leak basally into the hemolymph.

We added to our new revised version of the manuscript a new transgenic fly using a Dpp cDNA. Here GFP is inserted into the Dpp coding sequence after the site of Furin cleavage in the Golgi, which separates the Dpp propeptide with its own signal peptide and the C-terminal rest of the protein. In this construct, we engineered a STOP codon at the GFP C-terminus so that the mature peptide is not translated (see scheme in NEW Extended Data Figure 13d). Therefore, upon Furin cleavage of this chimeric protein, only GFP is produced without the C-terminal mature peptide which is not translated. This mature C-terminal fragment is the morphogen that generates a gradient. Using the Gal4 system, we express this construct by means of a Dpp enhancer region. Therefore, this molecule (which we term sGFP^{Dpp}) retains the 5'UTR of Dpp, the signal peptide of Dpp and the cleavage sites of Dpp followed by GFP: GFP is transcribed, translated, cleaved and trafficked like Dpp. Note that there is no receptor that retains GFP at the membrane of the epithelial target cells: in this system there is only diffusion and, potentially, leakage. We have therefore generated a simplified system to study a potential role of leakage using a GFP molecule which shares with Dpp all the features in the producing cells, including its secretion properties. In summary, this is what we saw with this reagent:

a) Unlike sGFP$^{Ptc\text{-}Wg}$, in the case of sGFP^{Dpp}, GFP is robustly seen in the target tissue. Most sGFP^{Dpp} is seen in a 1µm thick region within the epithelium in the apical side (NEW Extended Data Figure 13c). This implies that, unlike sGFP$^{Ptc\text{-}Wg}$, sGFP^{Dpp} is secreted such that it does not leak out from the epithelium. It is possible that the locale of secretion is behind the difference between the two molecules.

b) In the target epithelium, this sGFP^{Dpp} does not form a gradient (see NEW Extended Data Figure 13e,g), indicating that the actual levels of leakage and diffusion in the wing cannot generate a gradient by itself. If we consider the diffusion coefficient of a molecule of the size of Dpp in the aqueous environment of the extracellular space without a binding partner ($100\ \mu m^2/s$, using Stokes-Einstein relation), we obtain an upper bound for the leakage rate k_L of $0.0005\ s^{-1}$ (assuming a potential decay length of Dpp about 3 fold the target tissue length

L). These estimations are conservative since the diffusion of GFP in the extracellular space is likely to be lower than in the aqueous environment, and we observe no decay in the GFP profile suggesting that the GFP decay length is more than 3-fold larger than the target tissue size. Note that this upper bound for the leakage rate k_L obtained in our measurements, is 150-fold smaller than $k_L = 0.075\ s^{-1}$ estimated in the JPV paper for sGFP$^{ptc\text{-}Wg}$, confirming that sGFP^{Dpp} has different leakage properties, and in particular, that in the Dpp morphogen system leakage is very small if at all existing.

c) In the JPV paper, the authors estimate that the leakage rate is 0.075 s^{-1}, corresponding to a time scale for the disappearance of molecules from the epithelium of 13 seconds. If this was true for sGFP^{Dpp}, then one would expect that sGFP^{Dpp} disappears from the epithelium in the timescale of 13 seconds. To address this, we cultured wing discs *ex vivo* at physiological temperature (25˚C) as requested by this reviewer and measured whether the sGFP^{Dpp} signal decayed after one hour (see NEW Extended Data Figure 13g,h). We could not detect any decay in the signal, confirming that sGFP^{Dpp} does not leak in one hour time *i.e.* well beyond the 13 seconds expected for the estimation of the leakage rate in the JPV paper. This is indeed consistent with the fact that sGFP^{Dpp} does not form a gradient. We repeated this experiment at 4˚C, to exclude the influence of *de novo* production of sGFP^{Dpp} in this experiment.

iii) We have determined the Dpp transport parameters incorporating leakage and assuming a leakage rates between 0.00001 and $0.001\ s^{-1}$ (NEW Extended Data Figure 13i,j) as prompted by our experiments with sGFP^{Dpp}. The trafficking parameters we obtained and the contribution of the different modules under these conditions remained unchanged (see Supplementary Material Section 2 chapter 3.10).

- Lack of knowledge of the molecular identity of the putative internalization receptor implicated by the analysis is a potential weak spot. If this receptor was identified, manipulation of its expression would provide a convincing test of the authors' model. Moreover, the apparent absence of evidence of what this receptor might be is troublesome. Given the intensity with which this system has been studied, candidates from genetic screens might be expected because mutations would be predicted to change the Dpp gradient and therefore wing patterning.

We proposed in the manuscript that the internalizing receptor(s) could be Dally, possibly together with Dally-like (Dlp).

page 8: *"In addition, Dally binds Dpp[3] and it can act on Dpp by facilitating its diffusion in the extracellular matrix milieu[4]. This could explain how Pentagone/Dally influence D_0. Furthermore, Dally is internalized[5-10] and can therefore act as one of the internalization receptors which mediates Dpp endocytosis. We speculate that, if Dally does function as one of the receptors which internalize Dpp, then the k_{on} and k_{off} rates that are regulated by Pentagone include the binding/unbinding rates of*

Dpp to Dally (for further discussion on how Pentagone/Dally could modulate D_0, k_{on} and k_{off}, see SM 4.5.2-4.5.4) ».

***page 66 in Supplementary Material Section 2 chapter 4.5.4**: " The rates k_{on} and k_{off} describe the binding properties of the ligand to the receptor that mediates its internalization. GPI-anchor proteins can be internalized[5-10] and can therefore act as one of the internalization receptors which endocytose Dpp. One can speculate that, if Dally does function as one of the receptors which internalize Dpp, then the k_{on} and k_{off} rates that are regulated by Pentagone include the binding/unbinding rates of Dpp to Dally.*

However, in the mutants for Pentagone or Dally, Dpp is still internalized. This implies that other receptors also contribute to Dpp internalization. Interestingly, it has been shown that in a double mutant for the two HSPG in the Drosophila genome, Dally and Dally-like (Dlp), the trafficking of Dpp is dramatically affected (Figure 6 in [11]). »

In order to address the concern of this reviewer about the internalization receptor, in the new version of the manuscript we study whether Dally together with Dlp have an effect in Dpp endocytosis: we abolished both their function by downregulating Sulfateless expression in the posterior compartment through RNAi. Sulfateless is the enzyme that catalyzes the N-sulfation of the glucosamine in HSPGs, which is a prerequisite for heparane biosynthesis and therefore for HSPG production. We observed that Dpp internalization was dramatically affected upon Sulfateless silencing (see NEW Extended Data Figure 14g-i).

If Dally and Dlp act as Dpp internalization receptors, it must be essential that these proteins are associated to the plasma membrane. To address this, we cleaved Dally and Dlp away from the plasma membrane by incubation with the PI-PLC enzyme. Under these conditions, we observed a considerable reduction of the Dpp endocytosis rate (see Extended Data Figure 14e,f).

These results indicate that together, Dally and Dlp might play a role as Dpp internalization receptors. Consistently, both *dally* and *sulfateless* mutants show a strong effect on the range of the Dpp gradient (Extended Data Figures 14g,h and 6b,d; see also reference[12]).

In the new version of the manuscript we therefore added:

***page 8**: "Consistently, cleaving HSPGs from the membrane by PI-PLC enzyme and abolishing its function by downregulating Sulfateless expression by RNAi dramatically affected the internalization of GFP-Dpp and the range of the gradient (Extended Data Figs. 14e-i and 6b,d; see also reference[12]).*"

we also covered this in Supplementary Material Section 2 chapter 4.5.4

page 67: "In order to study whether Dally together with Dlp have an effect in Dpp endocytosis, we abolished both their function by downregulating Sulfateless expression in the posterior compartment through RNAi. Sulfateless is the enzyme that catalyzes the N-sulfation of the glucosamine in HSPGs, which is a prerequisite for heparane biosynthesis and therefore for HSPG production. We observed that Dpp internalization was dramatically affected upon Sulfateless silencing (see Extended Data Figure 14g-i).

If Dally and Dlp act as Dpp internalization receptors, it must be essential that these proteins are associated to the plasma membrane. To address this, we cleaved Dally and Dlp away from the plasma membrane by incubation with the PI-PLC enzyme. Under these conditions, we observed a considerable reduction of the Dpp endocytosis rate (see Extended Data Figure 14e,f).

These results indicate that together, Dally and Dlp might play a role as Dpp internalization receptors. Consistently, both dally and sulfateless mutants show a strong effect on the range of the Dpp gradient (Extended Data Figures 14g,h and 6b,d; see also reference[12])"

Referee #3 (Remarks to the Author):

I appreciate the thorough and well-thought responses of authors to my concerns. I am fully satisfied with their arguments and with the additional data provided to support these. A small but very important issue, though. I wonder why the transcytosis module in their original equation. There is no convincing experimental evidence supporting this mode of Dpp movement in this particular model system. Could authors briefly argue why? Is this module absolutely necessary? Do their data fit removing it? If so, why not doing so?

We thank the referee for the assessment of our revised manuscript and are pleased that he/she is fully satisfied with our arguments and additional data.

Regarding the transcytosis module, there was no reason *a priori* to exclude transcytosis as potentially contributing to the transport of Dpp. It was only after careful evaluation of the data that we reached the conclusion that the transcytosis module has minimal impact on Dpp gradient formation. This also highlights that our physical description can capture all possible models; including those that, eventually, are not relevant to our system.

Although the transcytosis module is essentially irrelevant to Dpp gradient formation in this specific model system, removing it all together from our physical description is highly nontrivial. This is because this module was not explicitly included by construction, but is instead an emerging property of the cell-based description of the system. To directly exclude transcytosis one would need to write a more detailed set of equations to describe the system that would require additional

parameters and assumptions (*e.g.* the width of the intercellular space etc). A more simple and general approach is therefore more appropriate and does not impact our conclusions.

Referee 4

The autors have revised their manuscript by providing more data as well as clarifications.

We thank the referee for his/her careful assessment of our MS.

Revisions, however, did not significantly change some critical aspects of this study (see points 1-4 below).

We are confident that these remaining concerns are mainly due to misunderstandings, and we provide clarifications and additional data to resolve these issues below and in the revised manuscript.

I also agree with referee #1 when she/he writes that "it is difficult to assess the rigorousness of the experiments and the confidence one should have in the conclusions" (referee #1). This somewhat echoes some of the points that I raised.

The comment of referee 1 mentioned here by referee 4 is pertinent to the first round of revisions. We have already addressed this concern in our previous version of the manuscript and referee #1 has not raised any further issues with regards to this point.

Otherwise, a recent study of an engineered GFP gradient from the lab of JP Vincent, cited as ref #4 in the manuscript, showed that diffusible GFP can replace Dpp in the presence of Dpp receptors rewired to bind GFP. This suggests that regulated trafficking of the ligand may not be critical to form a scaling gradient. The authors further showed that "concomitant expression of glycosylphosphatidylinositol (GPI)–anchored non-signaling receptors further improved patterning, to near wild-type quality. Theoretical arguments suggest that GPI anchorage could be important for these receptors to expand the gradient length scale while at the same time reducing leakage" (abstract). This paper is fully relevant for this study since it provides experimental evidence indicating that tissue-level diffusion of Dpp and GPI-anchored non-signaling binders may be sufficient to extend the gradient with no need for specific regulation of Dpp trafficking.

First, our manuscript and the paper from the JPV lab mentioned here do not contradict each other: they are orthogonal to each other. The JPV paper shows how a gradient could in principle be engineered through a synthetic approach which is based on a secreted GFP, receptor binding and leakage, while our manuscript combines a plethora of assays to reveal the actual cellular mechanism by which a particular morphogen, Dpp, forms a gradient that scales (expands as the

tissue grows). With respect to our work, the JPV paper has the value of uncovering a potential role for leakage; we took this seriously and evaluated to what extent leakage is important for Dpp gradient formation (see below).

Second, it is indeed fascinating that just by using GFP and synthetic receptors/coreceptors based on a GFP nanobody, the authors in the JPV paper could provide Dpp signaling which retains essential properties of Dpp signaling itself. However, the claim of the reviewer that this signaling gradient can be achieved simply by diffusion, receptor binding and leakage without intracellular trafficking is not supported by data in the JPV paper.

Indeed, that paper consists of two parts. In the first part (their figures 1-3), the authors use secreted GFP which is detected as a gradient only when membrane bound nanobodies are present (no GFP gradient can be detected in the absence of those nanobodies); the authors observe an apical and a basal gradient with different properties. The authors then go on to only study gradient formation of the basal part only. There is an offset in this basal gradient, which the authors interpret as a signature of leakage back from the hemolymph to the epithelium. However, the basal gradient only represents 10% of the total signal: the vast majority of the signal is apical. Note that no sign of leakage or offset is seen in the apical gradient.

The second part of the report (their figures 4,5) is where signaling is studied which is indeed fascinating. In this second part the authors generate a Dpp signaling gradient synthetically based on GFP and an anti-GFP nanobody receptor which contains in this case a cytosolic tail taken from actual Dpp receptors, Thickveins and Punt. Here the authors study signaling, but they cannot study gradient formation of the synthetic ligand. This is for three reasons:

i) In this scenario, the GFP gradient is not detectable (see Figure S10d-f in the JPV paper; "…note that no GFP gradient was detectable in this genetic background…").. Therefore, nothing can be said about the transport of these GFP molecules, the existence or not of offsets, the existence or not of leakage, or whether molecules were internalized or not by endocytosis.

ii) If there is leakage in these conditions, as in the system in the first part of the JPV paper, the leaking pool would only be basal and represent only 10% of the total gradient. Instead, signaling could be elicited by the apical, non-leaking pool (the 90%). Therefore it is not known, whether only diffusion and leakage could generate the observed signaling gradient. The authors simply do not see the GFP gradient in the second part of the article.

iii) The cytosolic tails of Thickveins and Punt in the rewired nanobody receptor are known to contain domains which elicit endocytosis and recycling. Therefore, endocytosis and recycling are expected to be mediated by these receptors and thus could be important for the intracellular trafficking of the GFP in the second part of the paper (unlike the membrane bound nanobodies in the first part which lack those Thickveins/Punt domains).

In summary, the synthetic signaling engineered in the JPV article is fascinating, but it is not connected to the formation of the ligand gradient. Also, the JPV paper is interesting in that it explores what is possible, while our paper studies what does happen for Dpp by making a tour de force effort to establish assays and measure the most important parameters in the system. What we discovered is that these are the key features of Dpp gradient formation: i) recycling of Dpp takes place in the wing, ii) recycling is essential for gradient formation and iii) the contribution of recycling as the disc grows becomes increasingly important, explaining gradient scaling. Prompted

by the JPV paper, we investigated the potential role of leakage, and we found that it is not important in our system (see below).

Despite its relevance, this study is first cited to indicate that "unbound ligand molecules could in principle leak out of the epithelium". This is not appropriate.

In the revised version of our manuscript, the JPV paper appears 3 times in the main text in the context of leakage, GPI anchor non-signaling receptors, and the extracellular-diffusion model (pages 3, 5 and 8). It is also the center of chapters 2.11, 3.10 and 4.4 in Supplementary Material Section 2. We have performed new theoretical analysis to address a potential role of leakage, generated a new secreted GFP which is translated, cleaved and secreted like Dpp itself (not based on Wingless and Patched sequences as in the JPV paper) and performed chase experiments with this reagent to address important issues raised in the JPV paper. We hope that the reviewer finds the references to the JPV paper in our manuscript appropriate now.

1. I maintain that the main claims of the paper (scaling is achieved by increased recycling of Dpp, driven by Pentagone-dependent changes in the binding rates of Dpp to its internalization receptor, see abstract) are only supported by indirect evidence. As indicated earlier, the internalization receptor of Dpp is unknown and binding rates can only be indirectly inferred.

We respectfully disagree that our conclusions are "only supported by indirect evidence". In this manuscript, we make a "tour the force" effort to support each and every statement we make by quantitative measurements for which we also estimate uncertainties.

About the criticism concerning the internalizing receptor raised here, we propose that Dally is one of the internalization receptors for Dpp in our manuscript. We estimated the binding/unbinding rates to the receptor that internalizes Dpp by a combination of five independent assays, which strongly constrains the possible values for these rates (Extended Data Figure 6e,f). In our approach estimating these values, we systematically determined the confidence intervals. Using PI-PLC incubation and *sulfateless* mutant and studying their impart on Dpp endocytosis, we have now showed that indeed Dally and Dlp behave as internalization receptors for Dpp (NEW Extended Data Figure 14e-i).

Note also that all measurements in science are the result of indirect inference. Let us consider two examples. First, when the mass of an object is measured with a balance, what we measure are the deformations of a spring from which one can infer the force exerted from a gravitational field. Second, the standard way of measuring binding rates is surface plasmon resonance, where plasmon resonance changes with the binding states of molecules attached to a surface. The resulting change in the angle of minimum light reflection is used to infer the binding rates. The criticism that our measurements are indirect becomes relative in the context of these considerations.

2. The authors have taken into consideration my comment about the photoconversion of internalized Dpp and now acknowledge that this assay addresses whether internalized Dpp CAN contribute to gradient formation. In this revised version, the authors provide evidence that Dpp CAN be re-exocytosed from the target tissue (Fig 1b, new panel).

This experiment appears at the beginning of the manuscript, shows that internalised Dpp CAN be re-exocytosed and motivates our approach to measure all trafficking rates to assess what is the relevant pathway for Dpp transport, whether recycling is indeed relevant to how far reaches the gradient. We conclude that the pathway reflected by the recycling MODULE is relevant in early and late discs. This reviewer admits that "...In this revised version, the authors provide evidence that Dpp CAN be re-exocytosed from the target tissue (Fig 1b, new panel)…", which is indeed what we intended to show with that first experiment in the text (the one that motivates the rest of the analysis), the photoconversion experiment.

Of note, i) no quantification is provided

We would like to point out that this experiment was indeed quantified already in our previous version (see Figure 1c, Extended Data Figure 5h).

and ii) formation of the gradient is observed over the source.

The photoconversion experiment is used in the paper to show that Dpp can be re-exocytosed and that this re-exocytosed Dpp molecules can make a gradient. In the first version of our manuscript, combining nanobody uptake and acid wash, we photoconverted only the internalized pool of GFP-Dpp/GBP-Dendra2. Photoconversion was performed in the source region (where concentration is highest). In fact, which is the precise photoconversion region is in principle irrelevant, because only endocytosed molecules are converted. We saw that the photoconverted molecules appeared in the target region outside the source: internalized molecules can be re-secreted and travel in the target area.

Reviewer comments relevant to the first version of our manuscript, made us consider the possibility that the source has special properties: the endocytic and secretory pathways could have mixed in the source and piggybag endocytic cargo to be secreted this way only in the source area.

In the revised version of our manuscript, we photoconverted the internalized GFP-Dpp/GBP-Dendra2 in the target area. From there, these molecules spread into the source. The fact that molecules leave the target area means that those internalized molecules can be re-secreted from the target. Therefore the source is not special, the target cells can also re-secrete internalized molecules.

However, the extent to which internalized Dpp DOES contribute to gradient formation remains unaddressed. Despite this, the authors did not change their conclusion ("therefore recycling contributes to gradient formation"). Thus, the issue remains.

The referee seems to suggest that we do not know what the contribution of recycling is. This is a surprising comment as the core of the paper is about quantifying the degree to which recycling is relevant. Not just to know whether there is Dpp recycling, but to quantify the contribution of recycling to how far Dpp molecules move in the tissue (how large is the decay length of the Dpp gradient). This is the reason why we introduced a new concept: the Modules. The rational and meaning of this new Module concept is explained in Supplementary Material Section 2 chapter 3.1 and page 5 of the main text. We conclude how important the recycling MODULE is: the recycling module accounts for about 90% of the decay length (λ^2) in late discs, and about 50% in early discs. The recycling module does not contribute to the decay length in Dally and Pentagone mutants.

We tried to falsify the existence and relevance of recycling. We did this using five additional independent experiments (see below). In particular, to further investigate the role of recycling in gradient formation, we look at the effect of downregulation of recycling on the Dpp gradient. In order to downregulate recycling, we have silenced the expression of the recycling Rab proteins: Rab11 and Rab4. We performed RNAi experiments in which we express validated dsRNAs for Rab11 or Rab4 in the posterior compartment of the wing disc. We have observed a decrease of the decay length of eGFP-Dpp^{LOP} gradient (see NEW Extended Data Figure 12a-c), indicating that recycling plays an important role in gradient formation.

In view of all these facts, we believe that the comment of this reviewer that, *"…the extent to which internalized Dpp does contribute to gradient formation remains unaddressed…"* is a reflection of a misunderstanding. We hope that the clarifications we provide here and in the revised manuscript have helped resolve this.

Furthermore, in the revised version of the manuscript, we have very significantly strengthened the data on blocking recycling Rab proteins. First, we have measured the recycling rate of Dpp by the nanobody assay upon silencing this Rab proteins and in these conditions the recycling rate of Dpp is significantly decreased. We then checked how the decay length of the Dpp gradient is affected under these conditions. We had already performed this kind of analysis in the previous version of the manuscript, but in the current version we have increased the sample size and performed a thorough statistical analysis (NEW Extended Data Figure 13). The data shows a significant decrease in the decay length of the gradient when downregulating recycling (NEW Extended Data Figure 13). Furthermore, when Rab4 is silenced, the eGFP-Dpp^{LOP} gradient does not scale: a mutant in endocytic recycling is a scaling mutant. This indicates that indeed recycling impacts the decay length of the gradient dramatically: recycling is important for gradient formation and scaling.

3. “The experimental data are not meant to challenge the model but to make it work”. The point I raised about the model and its parametrization echoes the first comment of referee #2 (“observations and conclusions are a consequence of the particular modeling assumptions”). The authors respond to our criticisms by stating that they “performed here a minimal description of phenomena that we know happen in the tissue” (p. 18) and that their “approach is not dependent on a particular model, but instead consists of developing a phenomenological description that takes into account processes that we know occur in the tissue” (p. 6). They further state that “the transport of Dpp with 5 pools and 8 rates is not a model, it is a physical description” and that “these effective pools and rates do exist”. These statements are somewhat puzzling and overall naive. This study provides a good example of forward modeling based upon known, or suspected, causalities (“processes that we know occur in the tissue: diffusion, binding etc…”) that are expressed in the form of a set of equations. As with any other model, the model proposed here is based on a set of assumptions. Some of these assumptions are clearly spelled out (gradient at steady-state, 5 pools of Dpp with various transport mechanism), others are hidden in the model (e.g. no leakage of Dpp into the extracellular space). Since the model studied here is derived from phenomenology, the model is essentially a test of its assumptions and must be falsifiable (on this topic, I encourage the authors to read the review by J Gunawardena https://doi.org/10.1186/1741-7007-12-29). In this study, the authors have not attempted to falsify their model but have parametrized a model to fit their experimental data.

The referee claims that our conclusions are a consequence of our assumptions. She/he suggests this by alluding to the fundamental truth that if a scientific hypothesis is not falsifiable, one does not obtain new information by testing it. However, we show in our work that we do obtain key novel information from the experiments that is not already contained in the model. In particular we identify the transport modules that dominate gradient formation and we reveal how the contributions of these transport modules change as the gradient scales.

To be concrete, our parameterization uses 5 completely independent assays to quantify the relative contribution of each module. Our analysis finds that the recycling module is important. We then attempt to falsify this model by using no less than FIVE other independent assays which were not used in the parameterization: i) Photoconversion, ii) iFRAP, iii) Dpp^{Timer}, iv) Tkv receptor clones, v) knocking down recycling Rabs (Rab4/Rab11). The photoconversion and iFRAP assays addressed whether there is recycling of internalized molecules (it could have been that while the parameterization yields a significant recycling, in these two experiments recycling was not detected); the Dpp^{Timer} assay addresses whether the age of spreading molecules in the extracellular space is compatible with slow phenomena such as intracellular trafficking and recycling (it could have been that the extracellular molecules were younger which would be incompatible with the existence of recycling); the Tkv clones experiment addresses whether the signaling receptor Tkv is the receptor that internalizes Dpp, which would not be consistent with our model predictions considering its known unbinding rates; the Rab analysis addresses whether recycling Rabs affect the recycling rate which we measured for Dpp, and whether recycling contribute to gradient formation (it could have been that in these Dpp recycling mutants the decay length was not affected). We are not aware of many papers in the field where falsification has been taken as seriously as in our work.

It is worth noting that a similar theoretical framework has been used before by different groups (Lander[13], Bollenbach[14] etc): the processes captured in our physical descriptions at the molecular

level are those widely accepted to be in place in this system across the field. Furthermore, if any of the processes we implement are not relevant, e.g. if one of the pools of molecules we consider is not present in our system, this will be simply reflected *a posteriori* once the data is fitted: the size of that pool will simply be negligible. For example, the existence of an extracellular pool is confirmed by direct measurement and the presence of an immobile pool is reflected in our FRAP analysis (also identified in other reports[13,15]). Moreover, not all modules emerging from our theoretical framework need to hold any weight following quantification. Their weight (i.e. their relative contribution) is an emerging property. For example, the transcytosis module turns out to be negligible in all experimental conditions considered in our analysis. Similarly, the unbound module carries very little weight in large discs but plays a role in small and Pentagone mutant discs. The recycling module does not play any role in gradient formation in Pentagone and Dally mutants. This shows, once again, that our conclusions are by no means "a result of our modeling assumptions".

In the first version of our manuscript, we did not consider leakage. When the JPV paper appeared after our initial submission, we performed experiments to evaluate its potential relevance and developed theory to study how leakage would affect the gradient and our parameterization of the transport rates. In this case leakage turned out not to be relevant (see Supplementary Material Section 2 chapters 2.11, 3.10, 4.3).

Finally, we are aware of the Gunawardena review mentioned here. In our opinion, what Gunawardena describes in this article as the paradigmatic way in which modeling and quantitative biology should be performed matches perfectly with what we did in the present work (e.g. "…Keep the complexity of the assumptions in line with the experimental context and try to find the right abstractions…"; "…it is standard practice to fit to a training data set and to falsify the model, once it is fitted, on whether or not it predicts what is important…" *etc*).

4. Precision of the measurements

While the authors agree that precise measurements are important for the "robustness of their final conclusion", they actually refer to SM sections 2.5.2 (that discusses how the confidence intervals were established for k_N, k_r and k_o) and 2.3.7 (that examines how relaxing threshold values for the calculated extracellular pool and for the fitting of the FRAP data) to indicate that their conclusions are robust to changes in the corresponding measured values. This may be seen as paradoxical. I still find difficult to evaluate how precise and reproducible are the results of the transport assays developed in this study.

We have used standard but sophisticated techniques to analyze the data in order to obtain not only estimates of parameter values, but to also obtain strong information about the reliability of the results. We therefore know which conclusions are strong and robust and which are weaker. An important example was the use of our experimental data on the extracellular pool when estimating trafficking parameters. From the experiments, we obtain a range of possible values for the size of the extracellular pool, which defined the parameter regime. We tested whether our results are

robust even if widening this uncertainty range to make sure that our results are not sensitive to our error estimates. We could show that our results exclude certain parameter regimes even if we widen (relax) the allowed range of the extracellular pool. This is not paradoxical, but stems from the fact that possible parameter regimes consistent with FRAP data were located in separate "clouds" in the k_{on}-k_{off} space. When we considered a larger uncertainty for measured value of the extracellular pool, the lower cloud of data from the FRAP analysis was still excluded. This shows that the measured values lie in the upper cloud (*i.e.* recycling is important) and that this result is robust to our error estimates. We have explained our parameter estimation and quality assessment in detail in the supplement (Supplementary Material Section 2 chapter 3.7).

5. Pentagone levels over time

The authors used a pent-lacZ transgene to examine the expression of Pentagone. This approach is sub-optimal as it only reports upon the activity of a 2 kb fragment of regulatory DNA, whereas FISH is an easy and quantitative approach to detect endogenous gene expression. Otherwise, it is not clear how discs area (Extended Data Fig 15b) relates to the l=80 and l=144 stages examined in the rest of the paper. Based on the data plotted in Ext Data Fig 11e, it appears that not a single disc in Extended Data Fig 15b corresponds to l=144. Overall, these data are inconclusive.

We have now quantified the levels of Pentagone expression by using a MiMIC fly line where the endogenous gene is tagged by GFP to study the endogenous levels. Individuals homozygous for the tagged gene are perfectly viable and without phenotypes. The conclusion of this alternative measurement confirms what we described for the pent-lacZ transgene: Pentagone expression increases as the wing grows. Note in particular a very significant increase of PentGFP from discs of posterior length of 80 µm to 144 µm, the standard lengths used in our report. In this new data set there is therefore no longer an issue with the range of sizes analysed (NEW Extended Data Figure 15a,b).

6. model

I must admit that I do not understand the authors response: does recycling increase at the source? If so, how does production rate remain constant despite an increase in recycling? If not, what kind of mechanism might prevent an increase in Dpp recycling only at the source?

There appears to be a misunderstanding here regarding the definition of production rate used in our work. The operational definition of production is the secretion of molecules which were not previously endocytosed. Thus, secretion involves a pathway from the ER/Golgi to the plasma membrane without involving recycling endosomes. This pathway is captured in the set of equations as a production rate (see equation S1.1 in Supplementary Material Section 1). In the dynamic equations, which capture the dynamics of Dpp (equations S1.1-S1.5), the production rate corresponds to the parameter ν, while the recycling rate is k_r. These are two

independent parameters. A change in one, does not necessarily cause a change in the other one.

It is worth noting that our data shows that the recycling RATE (k_r) does not change during scaling or in the mutants considered (*pent* and *dally*). It is only the recycling MODULE ($\lambda_r^2 = \frac{D_0}{k_{on}} \frac{k_{off}}{k} \frac{k_r}{k_i+k_1}$) that changes under these conditions. The recycling rate (green) is only part of what is important for the recycling module. Indeed, the recycling rate, is the rate that characterizes the return of endocytosed molecules to the extracellular space. Instead, the recycling module is an emerging property or process in the system that reflects the contribution of a subset of phenomena (including internalization, intracellular degradation and recycling, but also extracellular diffusion, binding and unbinding to receptors) to the decay length of the morphogen.

The production rate of Dpp has been measured previously[16] and it remains constant as the disc grows. The mechanism by which Dpp production rate remains constant is not in the scope of this paper.

7. Dpp overexpression
The authors have addressed my concern and now show that GFP-Dpp produced at the locus (CRIPSR) scale and recycle similarly as GFP-Dpp (LOP) (Ext Fig 14). Also, the authors now suggest that over-expression of Dpp has no significant effect on the behavior of GFP-Dpp (LOP). Note, however, that the relative levels of Dpp associated with these two genotypes were not determined.

The blot was quantified (see legend of Extended Data Figure 7a). We made this more visible in the main test of the revised version.

page 3: "We used the LexA/LOP system to express eGFP-tagged Dpp at quasi-endogenous levels (eGFP-Dpp^{LOP}, see [17] and Materials and Methods and Western Blot analysis and quantification in Extended Data Fig. 7a). »

8. Dpp recycling and photoconversion See point #2 above.

9. scaling and data selection
- Fig 2a was revised
- raw data are now plotted in Ext Fig 6a, showing the linear regression

We are pleased that these points are clarified.

10. nanobody internalization assay/iFRAP/transcytosis/Dpp timer/Shape of the gradient in Pentagone mutants
The authors have clarified the issues raised in my review

We are pleased that these points are clarified.

I also provide below a few other comments about issues raised by referees #1-3

Please note, these are the issues concerning the previous version of the manuscript which were addressed and satisfied both reviewer #1 and #3. See details below.

referee #1
- p. 3, Fig 1a,b: the authors did not address the point raised by the referee and maintains their conclusion: "photoconverted molecules were ONLY in the endocytic pathway" (also Refree #2, p.11).

We did address this point and our response has satisfied referee #1. We also explained again this same point for this reviewer (#4) in our response here (see above).

- p.3-4, leakage into the extracellular space: the absence of leakage is one assumption of the model. The authors argue that the formation of the Dpp gradient is unlikely to be influenced by the loss of molecules in the extracellular space based on a recent paper by J.-P. Vincent and colleagues (2020) who actually showed the opposite for extracellular GFP: "in conclusion, a single extracellular binding species reveals the gradient of an inert protein in vivo, but leakage in the hemolymph occurs and interferes with the gradient's shape, most obviously far from the source, at the tail end of the gradient." The critical point made by the authors in their rebuttal is that leakage is an issue only for the ligands that are secreted baso-laterally. Since signaling Dpp is apical, leakage is not considered to be a significant issue: "GFP-Dpp is present in the apical-most 5mic of the epithelium, where there is no leakage, we did not consider leakage as a major phenomenon in the system." While I understand the logic of this argument, whether GFP-Dpp leaks into the extracellular space remains an assumption of the model that is not sufficiently well tested. In Ext Data Fig 13, the fraction of GBP-Alexa555 leaking into the extracellular fraction is not measured and no positive control is included.

The criticism here appears to be that we have not sufficiently tested whether "…GFP-Dpp leaks *into the extracellular space*…" and that "the fraction of GBP-Alexa5554 leaking *into the extracellular fraction* is not measured" (the emphasis is ours). We address this criticism considering three points:

i) the reviewer requests that we should measure leakage from the hemolymph back "into the extracellular fraction". This back-leakage is irrelevant to the potential role of leakage to Dpp gradient formation according to the model proposed in the JPV paper.

ii) we have generated a new reagent (sGFP^{Dpp}), where secreted GFP is transcribed, translated, cleaved in the Golgi by Furin and secreted like Dpp, unlike the Patched/Wingless-based secreted GFP in the JPV paper. This new sGFP^{Dpp} shows clearly that, when secreted like Dpp in the apical side of the epithelium, leakage of this molecule is negligible (NEW Extended Data Figure 13e-h; see Supplementary Material Section 2 chapter 2.11).

iii) we have performed theoretical analysis where now leakage is included in our 5 dynamic equations. This analysis shows that, with the leakage rate estimated for $sGFP^{Dpp}$, our conclusions are not affected: the changes in the values of the different transport rates caused by leakage are negligible and the importance of the recycling module remains unaffected (see Supplementary Material Section 2 chapter 3.10).

This data strengthens the conclusion of the previous version of our manuscript that leakage is not important for Dpp gradient formation. We discuss this in details below:

i) The reviewer requests that we should measure leakage from the hemolymph back "into the extracellular fraction". In the JPV paper, the authors take the existence of an offset in the sGFP gradient as an indication that not only sGFP molecules leak out of the epithelium into the hemolymph, but they also can "leak" back from the hemolymph *into the extracellular space* to generate an offset in the gradient profile. Because the authors realize that there is no prominent offset in morphogen gradients such as Dpp, they propose that some organ(s) in the larva serve as sink which removes the morphogen from the hemolymph and prevents its return back to the wing as an offset. To mimic this sink, the authors express the anti-GFP nanobody receptor in the fat body which traps sGFP from the hemolymph and leads to the disappearance of the gradient offset. Therefore, because the Dpp gradient does not show a prominent offset, the leakage back from the hemolymph to the wing is not expected to happen.

ii) It is well established, that in the wing epithelial cells, different membrane proteins (HSPGs vs. Notch vs. EGF, etc) traffic in the secretory pathway through different, non-overlapping populations of ER/Golgi organelles[1]. The targeting of these proteins to apical versus basal secretory compartments affects the location (apical vs basolateral) in which those molecules are secreted. The specific targeting of mRNA to a subset of ER/Golgi compartments is encoded in non-translated 5' and 3' UTR sequences in the mRNA as well as on signal peptide sequences at the protein level[2].

This has in particular been thoroughly studied for the case of Wingless.

The reagent in the recent study by Stapornwongkul et al.[18] was generated by using the 5'UTR of *patched* fused to the coding region of the signal peptide of Wingless followed by the coding sequence of GFP downstream (we term this here $sGFP^{ptc\text{-}Wg}$). This construct is inserted in the endogenous *patched* gene and is therefore transcribed like *patched* (see scheme in Extended Data Figure 13c). The authors barely see $sGFP^{ptc\text{-}Wg}$ in the epithelium from what they conclude that these molecules are secreted basolaterally and leak basally into the hemolymph[18].

We have generated a transgenic fly using a Dpp cDNA where GFP is inserted into the Dpp coding sequence after the site of Furin cleavage in the Golgi, which separates the Dpp propeptide with its signal peptide (which remains in the Golgi apparatus of producing cells) and the C-terminal rest of the protein (the secreted ligand). In this construct, we engineered a STOP codon at the GFP C-terminus so that the mature peptide is not translated. Upon Furin cleavage, only GFP is produced

(see scheme in Extended Data Figure 13d). Using the Gal4 system, we express it by means of a Dpp enhancer region. Therefore, this molecule (sGFP^{Dpp}) retains the 5'UTR of Dpp, the signal peptide of Dpp and the cleavage sites of Dpp followed by GFP: the mRNA of GFP is transcribed and the protein, translated, cleaved and traffics like Dpp. Note that there is no receptor that retains GFP at the membrane of the epithelial target cells and that GFP is not endocytosed as it is only present in the extracellular space: in this system there is only extracellular diffusion and, potentially, leakage. We have therefore generated a simplified system to study a potential role of leakage using a GFP molecule which shares with Dpp all the features of production, trafficking and secretion characteristics of the producing cells. In summary, this is what we saw with this reagent:

a) Unlike sGFP$^{ptc\text{-}Wg}$, in the case of sGFP^{Dpp}, GFP is robustly seen in the target tissue. Most sGFP^{Dpp} is seen in a 1µm thick region within the epithelium in the apical side (NEW Extended Data Figure 13c). This implies that, unlike sGFP$^{ptc\text{-}Wg}$, sGFP^{Dpp} is secreted such that it does not leak out from the epithelium. It is possible that the locale of secretion is behind the difference between the two molecules.

b) In the target epithelium, this sGFP^{Dpp} does not form a gradient (see NEW Extended Data Figure 13e,g), indicating that the actual levels of leakage and diffusion in the wing cannot generate a gradient by itself. If we consider the diffusion coefficient of a molecule of the size of Dpp in the aqueous environment of the extracellular space without a binding partner ($100\ \mu m^2/s$, estimated using Stokes-Einstein relation), we obtain an upper bound for the leakage rate k_L of $0.0005\ s^{-1}$ (assuming a potential decay length of Dpp about 3 fold the target tissue length L). These estimations are conservative since the diffusion of GFP in the extracellular space is likely to be lower than in the aqueous environment, and we observe no decay in the GFP profile suggesting that the GFP decay length is more than 3-fold larger than the target tissue size. Note that this upper bound for the leakage rate k_L obtained in our measurements, is 150-fold smaller than $k_L = 0.075\ s^{-1}$ estimated in the JPV paper for sGFP$^{ptc\text{-}Wg}$, confirming that sGFP^{Dpp} has different leakage properties, and in particular, that in the Dpp morphogen system leakage is very small if at all existing.

c) In the JPV paper, they estimate that the leakage rate is 0.075 s^{-1}, corresponding to a time scale for the disappearance of molecules from the epithelium of 13 seconds. If this was true for sGFP^{Dpp}, then one would expect that sGFP^{Dpp} disappears from the epithelium in this 13 seconds. To address this, we cultured wing discs *ex vivo* at physiological temperature (25˚C) as requested by this reviewer and measured whether the sGFP^{Dpp} signal decayed after one hour (see NEW Extended Data Figure 13g,h). We could not detect any decay in the signal, confirming that sGFP^{Dpp} does not leak in one hour time *i.e.* well beyond the 13 seconds expected for the estimation of the leakage rate in the JPV paper. This is indeed consistent with the fact that sGFP^{Dpp} does not form a gradient. We repeated this experiment at 4˚C, to exclude the influence of *de novo* production of sGFP^{Dpp} in this experiment.

iii) We have determined the Dpp transport parameters incorporating leakage and assuming a leakage rates between 0.00001 and $0.001\ s^{-1}$ (NEW Extended Data Figure 13i,j) as prompted by our experiments with sGFP^{Dpp}. The trafficking parameters we obtained and the contribution of the different modules under these conditions remained unchanged (see Supplementary Material Section 2 chapter 3.10).

referee #2

I agree with the referee when she/he writes that “most rates can be measured only indirectly” and that “inferred values are model-dependent”. This echoes my own evaluation (“the rate of Dpp recycling was not measured”, “the rates of Dpp binding to its receptors were not directly measured”, “analysis of the binding rates of Dpp to its putative internalization receptor cannot be performed”, “because most of these transport values (Ext table 1) are not experimentally accessible” etc…). As indicated above, the idea that this study provides a phenomenological description that is not dependent on a particular model seems to be naive. I am not sure to understand why the authors insist that their conclusion is not the consequence of modeling assumptions.

We have discussed this above (answer to point #3)

One key assumption discussed by the referee is that the gradient is at steady-state. The authors now provide FRAP data to indicate that fluorescence recovery (by Dpp transport) is much faster than growth. Specifically, in Ext Data Fig 11i, the decay length of the total pool measured prior to bleaching appeared to be very similar to the decay length measured after 30 min of recovery (variations in decay length presumably reflect variations in disc size). Thus, the assumption of a steady-state gradient is a reasonable one. Of note, the authors state (p. 4): “this justifies the quasi-state assumption for the decay length underlying the parametrization”. Again, it would probably be more correct to conclude on an assumption of the model.

We are pleased that the reviewer appreciated the work we performed to support the steady state assumption following the concerns raised by another reviewer, and that he/she finds that our steady state assumption is reasonable.

1 Yano, H. *et al.* Distinct functional units of the Golgi complex in Drosophila cells. *Proc Natl Acad Sci U S A* **102**, 13467-13472, doi:10.1073/pnas.0506681102 (2005).

2 Simmonds, A. J., dosSantos, G., Livne-Bar, I. & Krause, H. M. Apical localization of wingless transcripts is required for wingless signaling. *Cell* **105**, 197-207, doi:10.1016/s0092-8674(01)00311-7 (2001).

3 Akiyama, T. *et al.* Dally regulates Dpp morphogen gradient formation by stabilizing Dpp on the cell surface. *Dev Biol* **313**, 408-419, doi:10.1016/j.ydbio.2007.10.035 (2008).

4 Fujise, M. *et al.* Dally regulates Dpp morphogen gradient formation in the Drosophila wing. *Development* **130**, 1515-1522, doi:10.1242/dev.00379 (2003).

5 Guha, A., Sriram, V., Krishnan, K. S. & Mayor, S. Shibire mutations reveal distinct dynamin-independent and -dependent endocytic pathways in primary cultures of Drosophila hemocytes. *J Cell Sci* **116**, 3373-3386, doi:10.1242/jcs.00637 (2003).
6 Gupta, G. D. *et al.* Analysis of endocytic pathways in Drosophila cells reveals a conserved role for GBF1 in internalization via GEECs. *PLoS One* **4**, e6768, doi:10.1371/journal.pone.0006768 (2009).
7 Mayor, S., Parton, R. G. & Donaldson, J. G. Clathrin-independent pathways of endocytosis. *Cold Spring Harb Perspect Biol* **6**, doi:10.1101/cshperspect.a016758 (2014).
8 Sabharanjak, S., Sharma, P., Parton, R. G. & Mayor, S. GPI-anchored proteins are delivered to recycling endosomes via a distinct cdc42-regulated, clathrin-independent pinocytic pathway. *Dev Cell* **2**, 411-423, doi:10.1016/s1534-5807(02)00145-4 (2002).
9 Hemalatha, A., Prabhakara, C. & Mayor, S. Endocytosis of Wingless via a dynamin-independent pathway is necessary for signaling in Drosophila wing discs. *Proc Natl Acad Sci U S A* **113**, E6993-E7002, doi:10.1073/pnas.1610565113 (2016).
10 Norman, M., Vuilleumier, R., Springhorn, A., Gawlik, J. & Pyrowolakis, G. Pentagone internalises glypicans to fine-tune multiple signalling pathways. *Elife* **5**, doi:10.7554/eLife.13301 (2016).
11 Belenkaya, T. Y. *et al.* Drosophila Dpp morphogen movement is independent of dynamin-mediated endocytosis but regulated by the glypican members of heparan sulfate proteoglycans. *Cell* **119**, 231-244, doi:10.1016/j.cell.2004.09.031 (2004).
12 Yan, D. & Lin, X. Shaping morphogen gradients by proteoglycans. *Cold Spring Harb Perspect Biol* **1**, a002493, doi:10.1101/cshperspect.a002493 (2009).
13 Zhou, S. *et al.* Free extracellular diffusion creates the Dpp morphogen gradient of the Drosophila wing disc. *Curr Biol* **22**, 668-675, doi:10.1016/j.cub.2012.02.065 (2012).
14 Bollenbach, T., Kruse, K., Pantazis, P., Gonzalez-Gaitan, M. & Julicher, F. Morphogen transport in epithelia. *Phys Rev E Stat Nonlin Soft Matter Phys* **75**, 011901, doi:10.1103/PhysRevE.75.011901 (2007).
15 Kicheva, A. *et al.* Kinetics of morphogen gradient formation. *Science* **315**, 521-525, doi:10.1126/science.1135774 (2007).
16 Wartlick, O. *et al.* Dynamics of Dpp signaling and proliferation control. *Science* **331**, 1154-1159, doi:10.1126/science.1200037 (2011).
17 Harmansa, S., Hamaratoglu, F., Affolter, M. & Caussinus, E. Dpp spreading is required for medial but not for lateral wing disc growth. *Nature* **527**, 317-322, doi:10.1038/nature15712 (2015).
18 Stapornwongkul, K. S., de Gennes, M., Cocconi, L., Salbreux, G. & Vincent, J. P. Patterning and growth control in vivo by an engineered GFP gradient. *Science* **370**, 321-327, doi:10.1126/science.abb8205 (2020).

Reviewer Reports on the Second Revision:

Referee #1 (Remarks to the Author):

I am satisfied with the authors' response. The addition of the data addressing the potential of morphogen leakage from the wing disc strengthens the conclusions.

I remain unconvinced that the short format of the paper is the most appropriate for this type of study. There is ten times more material in the Supplemental information than the main body of the paper. Important and interesting results are buried and will probably be missed by most readers.

Referee #3 (Remarks to the Author):

I appreciate authors' fair response to my concern and, a such, glad to support publication of this ms.

Referee #4 (Remarks to the Author):

I thank the authors for addressing the points raised in my review. Regretfully, my reservations remain nearly in full.

About the possible role of diffusion in the formation of the Dpp gradient
The paper by JPV and colleagues (Stapornwongkul et al., 2020) shows that a functional gradient can be engineered by a synthetic ligand. This demonstrates that diffusion of Dpp may in principle be sufficient to produce a patterned wing. Obviously, a different mechanism may operate in nature and it is important and necessary to explore what is happening in vivo for endogenous Dpp. In this sense, the authors are fully correct to emphasize that this finding does not contradict their own conclusions. Having said this, the authors are not taking seriously enough the key finding from the JPV paper.
1. in their rebuttal, the authors insist that the synthetic signaling ligand engineered by Stapornwongkul et al. (2020) 'is not connected to the formation of the ligand gradient' and that 'with respect to our work, the JPV paper has the value of uncovering a potential role for leakage'. In my view, the Stapornwongkul et al., 2020 provide compelling evidence that diffusion may be sufficient to form a signaling gradient, raising in turn the possibility that simple diffusion of Dpp may likewise be sufficient. Are there clear data excluding this possibility? As explained in my earlier review, the claim by the authors that their modeling approach combined with experiments is totally unbiased, and therefore includes a model whereby diffusion only would operate, is not entirely convincing.
2. in the main text, the JPV paper is not cited for its main finding, i.e. a functional gradient can be engineered by a synthetic ligand, but rather to emphasize the importance of Dally in gradient formation (p.8, l. 296-8), which is quite misleading. Indeed, the JPV paper does not re-investigate the role of Dally in gradient formation but rather shows that GPI-anchored non-signaling binders, such as Dally for Dpp, contribute to the formation of the signaling gradient of the artificial ligand. Thus, the authors mis-cite this important paper in the field only to strengthen a relatively minor conclusion of their study. This is really not appropriate.

Indirect evidence (# 1)
I maintain that that the binding rates of Dpp to its internalization receptor have not been studied and note that the authors did not address this point in their rebuttal. The general discussion on measurements in science is interesting but slightly off-topic.
Otherwise, the authors now provide some evidence that Dally/Dlp modulate the endocytosis of Dpp. This observation is somewhat expected since these HSPGs are known to stabilize Dpp at the cell surface. So, while these data are consistent with the idea that Dally/Dlp act as Dpp internalization receptors, they fail to show that the binding rate of Dpp to its proposed internalization receptor varies in a Pentagon-dependent manner during scaling (abstract).

Role of Dpp recycling in gradient formation (# 2)
In their rebuttal, the authors are hammering that they have discovered that 'recycling (of Dpp) is essential for gradient formation' and that 'the contribution of recycling as the disc grows becomes

increasingly important'. As mentioned earlier, I respectfully disagree. Introducing the concepts of Modules does not address the extent to which internalized Dpp contribute to gradient formation. Otherwise, the Rab11/Rab4 data are interesting but they do not show that it is the recycling of Dpp that is important. In a diffusion model, gradient formation was shown to depend on the levels of signaling and non-signaling binders (Stapornwongkul et al., 2020). Thus, an alternative interpretation of the Rab11/Rab4 data is that trafficking defects of Tkv and/or Dally/Dlp are causing the Rab11/Rab4 mutant phenotypes.

Model (#3)
I find revealing that the authors highlight the characterization of the transport modules as key novel information on the system. My point remains in full.

Precisions of the measurements (#4)
I am not sure that the quality of the experimental data should be evaluated by their usefulness in defining a parameter range to be used in a specific model.

Pent (# 5)
the authors have addressed my specific concern. A new PentGFP line (using a MIMIC) was generated and used to show that PentGFP levels increase over time (Fig S15a,b).

Model (# 6)
I understand from the rebuttal that the production rate corresponds to the secretion of 'not previously endocytosed' Dpp by source cells, as estimated from FRAP recovery profiles in Wartlick et al. (2011). However, going back to the original paper, one reads that 'in FRAP, the "production rate" is actually a combination of transcription, translation and secretion, diffusion and degradation in the source, i.e. all the processes
that contribute to the flux of molecules out of the source' (pp. 40-41 of the SM in Wartlick et al., 2011). Taking the original definition, my question remains: assuming that Dpp can be internalized by source cells, how does production by source cells of Dpp diffusing into the tissue remains constant if recycling increases over time? If the authors instead assume that Dpp is not internalized by source cells, this assumption needs to be explained. In sum, my point remains.

Leakage (minor specific point)
The authors have now produced a new secreted GFP, referred to as sGFP(Dpp). In FigS13e, the authors expressed this molecule under the control of a dpp-Gal4 and showed that it accumulates apically far from its source, suggestive of both free diffusion and absence of leakage. These data are nice. Based on these, the authors propose that endogenous Dpp likewise does not leak. This is a surprising conclusion. As pointed out by referee #1 in her/his review, whether Dpp has the same characteristics as extracellular GFP remains an assumption. Again, I am not sure that one can conclude on the basis of these observations that 'the Dpp morphogen system leakage is very small if at all existing' (rebuttal). Otherwise, the observation that sGFP(Dpp) does not form a gradient is perhaps expected and not particularly meaningful given the absence of specific receptors for this artificial ligand (see Stapornwongkul et al., 2020).
Minor comment: the results shown in S13e would need to be better described: is this a fixed sample? What are the cells stained with phalloidin that are locate apically? If it is a cross-section along the a-p axis, why is the expression of sGFP(Dpp) discontinuous at the source? Also, the distribution of sGFP appear heterogeneous on the surface view (right panels), with both cell-to-cell variability and specific accumulation at cell-cell junctions: please comment.

Last, the authors 'focused on the target cells in the posterior compartment of the wing imaginal disc' (l. 58). A very recent paper from the Affolter lab (bioRxiv preprint doi: https://doi.org/10.1101/2020.11.23.394379) provides a rationale for this otherwise arbitrary choice. It would be useful for non-expert readers to cite this work.

Referee #5 (Remarks to the Author):

Review for "Morphogen gradient scaling by recycling of intracellular Dpp"

To give a quick context, I am writing this as a "fresh" Referee after the previous rounds of review and revision. I have read thoroughly the latest draft of the manuscript and the multiple rounds of referee reports and author rebuttals. After a brief introduction, I give my Referee report on the manuscript and also the rebuttals to points raised by referees.

The question of morphogen scaling has been a long-standing one and has been challenging to find clear mechanisms. Even in the apparently simple morphogen Bicoid, there is significant debate over scaling – including if it even happens. Therefore, the aims of this paper are laudable and likely to have a broad impact.

Overall, the central results of the paper are highly impressive. The range of quantitative experiments, combined with theory, are very powerful and help to provide evidence for a plausible mechanism of morphogen scaling. The genetics and controls generally seem performed to a high standard, giving confidence in the main conclusions.

Below, I discuss a number of the issues that have arisen during the review process and outline my thoughts regarding the author responses. I also raise a few further issues that should be addressed.

Major comments:

1. There is significant debate between the authors and the referees (particularly #2 and #4) about the modelling and what can be interpreted from this. I break this down below:

a. "Indirect measurements": Referees raise the concern that parameters are indirectly inferred and heavily reliant on the model choice. Relatedly, this then asks questions about the falsifiability of the model.

This is a valid concern, and it is certainly true that too many studies are not rigorous enough in how they incorporate model fitting and then how such parameter values are interpreted. A particular challenge in biology is that the range of mechanisms are so varied that it can be difficult to provide a definitive model.

Having said this, I am generally convinced by the Authors responses (with some important caveats below). The range of multiple independent assays to fit parameters is very powerful. Their statistical approach appears valid and provides a suitable level of uncertainty. Of course, their conclusions are model dependent, but they consider a broad model framework (though see (1.b)). Therefore, I am happy that the paper conclusions – e.g. the level of recycling and the role this plays in gradient scaling – are on a solid footing.

b. "The generality of the model": There is substantial debate through the previous rounds of revision about whether the model was being suitably general enough. The authors in the main paper state: "This description is equanimous: depending on parameter values, the same general description allows to capture a wide variety of transport models." Overall, with the below caveat, this is a fair statement and the results as presented in the main paper are generally convincing with regards to this.

However, in the response to reviewers, the Authors are somewhat bolder, e.g. response to Referee 2: "Therefore, our conclusions are not a consequence of particular modeling assumptions but are uniquely obtained in a single phenomenological framework". This is not entirely correct. As an example, as highlighted by Referee 4, the authors do not consider cytoneme transport (Roy et al

Science 2014). I believe in their original response they claim that cytoneme transport can be subsumed within the current model: e.g. "extracellular diffusion incorporates not just the Brownian motion of molecules, but also their binding and retention by the extracellular matrix". This is insufficient for cytoneme mediated transport, which cannot be described (at least in a simple manner) within the authors existing framework. For an example of modelling cytoneme transport modelling see: Improving the understanding of cytoneme-mediated morphogen gradients by in silico modeling, PLoS Computational Biology 2021.

Of course, cytoneme transport of Dpp in the wing disc remains contentious. Yet, it definitely merits mentioning as a possible oversight in the paper. I note that, to the best of my knowledge, the paper does not cite any cytoneme-mediated transport papers.

A related factor here is one of dimensionality. The Authors present a 1D spatial model. Yet, once one starts considering cytonemes (or other possible directed transport mechanisms) it is important to distinguish between the dimensionality of the transport mechanism; e.g. extracellular diffusion is effectively 3D whereas cytoneme transport is effectively 1D. Such considerations are currently not considered in the model. I don't expect new analysis at this point, but clearer discussion of limitations is merited.

In summary, I am happy with the Authors response to the valid critiques raised – particularly by Referee 2 and in both rounds of revision by Referee 4 - about model fitting and how such fitting can be interpreted. Their conclusions are as solid as can be, given current experimental limitations. However, I feel clearer caveats need to be included within the manuscript to make the above issues clearer.

2. Regarding specific issues raised by Referee #2:

a. Discussion of the Fried and Iber, Nat Comm 2014, I am satisfied by the Author response. The range of FRAP experiments are quite convincing with regards to the relative equilibrium state of the gradient.

b. Approximation given by Eq. 1. I am happy with the Authors response.

c. Improved pulse experiments and discussion. This has been adequately addressed (though see my further comments below).

d. The response to the Referee regarding the FRAP experiment analysis is generally convincing. However, the Authors state in their original response: "Dpp transport depends on eight parameters. Three of them (endocytosis, recycling, clearance and immobilization rate) are determined by means of the nanobody uptake assay (explained in Supplementary Material Section 2.2). A fourth rate, the degradation rate of the immobile pool, was determined by measuring the slow relaxation rate in a long-term FRAP experiment. The last three, the extracellular diffusion and the binding/unbinding rate were determined by means of the Monte Carlo sampling and Approximate Bayesian Computation (Supplementary Material Section 2.5.2)." Unless I'm missing something, this text in fact only refers to seven, not eight, parameters. I suspect this is a typo, but can the Authors clarify.

e. I fully agree with the Referee #2 that, given the large-scale extensive analysis, that all appropriate scripts should be made fully available. I hope the Authors generate a Github (or similarly accessible) repository for these codes. It would also be excellent if the authors also include the quantitative data sets. Then, other researchers have the freedom to test different analysis approaches or different models. This will substantially increase the impact of the paper and sets a clear standard for presentation of results in quantitative biology papers.

f. I agree that the Authors have satisfactorily dealt with the "small points" raised by the Referee

#2.

3. Related to a lot of the above discussion, Referee #4 highlights potential problems with the data interpretation and the subsequent interpretation. In particular: "However, the extent to which internalized Dpp DOES contribute to gradient formation remains unaddressed. Despite this, the authors did not change their conclusion ("therefore recycling contributes to gradient formation"). Thus, the issue remains." I understand the Referee point and think that such scepticism is justified. However, I am overall convinced by the Author response to this issue (given the caveat that I think the limitations in the approach needs to be made clearer in the main paper). I find the range of quantitative experiments genuinely impressive, and it's hard to see what more could be done given current experiment and imaging tools. Of course, future work using more sophisticated approaches may show the conclusions here are not exactly correct, but that's science.

4. There is significant discussion of the precision of parameter measurement. There are two major points here: (1) the Authors use sensible approaches and provide suitable ranges on parameters; (2) the description in the Methods, as highlighted by Referee #4 are currently insufficient. I did not find the description of sufficient depth (e.g. Supp Mat Part 2, 2.5.2). This would definitely benefit from publishing of the code alongside the paper, as currently the presented information is insufficient for another researcher to independently validate these results. Given that such output is central to the paper conclusions, such information (the code and the quantitative numbers used for fitting) should be made available.

5. Paper structure and format:

a. I found reading the paper very challenging. As highlighted by Referee #1, the paper is extremely dense and hard to parse at times. I appreciate the length limitation imposed by the journal, but I think that the paper itself can still be substantially improved in terms of readability. I suspect by this stage that the Authors are tired of editing it, but in its current form I don't think the paper will have the impact it deserves due to the very dense writing.

b. Related to the above point, the Supplementary Material Section II is not clear enough. To properly understand this paper, requires a reading of the Supplementary Material. Yet, this material is very dense. Notably, the page numbers of the index do not match the actual page numbers, making navigation through the document more difficult than necessary. As with the paper itself, this material needs to be gone through properly and brought up to standard.

c. The Extended Figures should be ordered sequentially as they appear in the main text. It was frustrating moving back and forth through a complex series of documents.

In summary on this point, I found, with a lot of hard reading (and having five different documents related to the paper open on my computer at the same time) that the paper was rewarding and has very interesting results. However, in its current form, it is likely to lose some impact due to a degree of impenetrability.

6. Tandem reporter results:

a. I was surprised reading the results related to the tandem reporter that there was no mention (that I found) of how background signal was accounted for. The background signal can make a big difference to the measured ratio. In Durrieu et al MSB 2018, which used the tandem reporter to explore Bicoid dynamics, they used linear unmixing to provide background estimates for each individual sample. Here, how was background accounted for in the tandem reporter and what effect does this have on the reported ratio?

b. This is a minor point, but the bead assay is not necessarily giving the precise ratio between the two fluorophores (though it is likely close approximation). When imaging live, there are additional

factors such as scattering through the tissue that affect the GFP and mKate2 channels differently. The ratio value F can be left as an additional fitting parameter. This would increase uncertainty in fitting parameters but I doubt this will make a large change.

7. Framing of the results: I appreciate that for a journal such as Nature that the results must be seen to have broad implication. It is important to note I think the results presented here are of sufficient interest and breadth to be suitable for Nature. Having said that, there are times in the paper where I feel some "overselling" occurs.

a. The abstract states: "Here we unravel the cell biological basis behind scaling…" . I feel a more accurate statement would be along the lines of "Using formation of the Dpp gradient in Drosophila we unravel a cell biological basis behind scaling." Essentially, at a minimum I ask that "the" in the original sentence be changed to "a".

b. In the abstract: "… which ultimately allows adaptation of shape and pattern to different sizes of organs in different species." This is too bold and needs to be more balanced. The initial part of the sentence does have "could", but this sentence currently reads as claiming more certainty than is there given the results presented are only for one morphogen in one species.

c. First intro paragraph: "To understand how the gradient decay length is regulated during scaling, it is essential to determine which cellular transport phenomena are important during Dpp gradient formation and which of these change as the tissue grows." This sentence jumps from a broad point to a specific case – that of Dpp. In its current wording, it could be misread to mean that by simply understanding Dpp gradient formation we can understand gradient formation of all morphogens, which is clearly incorrect. I suggest splitting this into two distinct sentences – one on the general question of understanding decay length scaling, and a second on using Dpp as a specific example to tackle this problem.

d. "Therefore, recycling contributes to gradient formation and could in principle underlie scaling." This sentence is too bold at this point, given the evidence that has been presented up to this point. I would suggest removing or reworking to be more balanced (nearly the same sentence at the end of the manuscript would be fine).

e. "In our experiment, the measured Dpp Timer age in the extracellular and intracellular space is the same". Given the experimental uncertainty (large error bar), it is not appropriate to say "is the same"; "are similar" feels more appropriate.

Minor comments

- Line 48 typo (missing "a") - "…such a gradient is formed."
- Paragraph starting line 48. Two sentences starting "Because…" . This reads as very clunky.
- Fig. 1d is referenced prior to Fig. 1c
- I appreciate space is an issue but Fig 2f and Fig. 2g should have y-axis labels even if stated in legend
- "Each of these assays depends on most transport rates. We therefore perform a sequence of steps using the assays, where each assay adds further constraints to the parameter values (SM 2.5.2)." I find this sentence too vague. At least give the central idea of the fitting, as it is essential to the paper conclusions.
- Line 134, should be Fig. 3a,b I think
- At the point that Fig. 3a is introduced, there is no mention of the different measurements in wings of different size. Therefore, when first reading it is very surprising to see "l=144" and "l=80" as it's not clear what this means. It would help to have a sentence before this stating that measurements were done at two general wing sizes.
- "Pentagone mediates this shift." This is a somewhat odd sentence – no references or data are presented to back up this statement.

- “We show here that scaling of Dpp gradients in dally mutant discs perturbed”, should be “We show here that scaling of Dpp gradients in dally mutant discs is perturbed”
- Figure 2 legend, “ℓ = 144.6 ± 4 μm” – the level of significance on the two numbers is not the same. I assume it should be ±4.0.

In conclusion, this is, as the Authors themselves state, a “tour de force” to tackle the mechanism for how Dpp scales. This provides important new insight into an essential process in development. It will be suitable for publication with some careful rewriting and cleaning up the last remaining issues highlighted above and by the other Referees.

Author Rebuttals to the Second Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

I am satisfied with the authors' response. The addition of the data addressing the potential of morphogen leakage from the wing disc strengthens the conclusions.

I remain unconvinced that the short format of the paper is the most appropriate for this type of study. There is ten times more material in the Supplemental information than the main body of the paper. Important and interesting results are buried and will probably be missed by most readers.

We are glad that the referee is satisfied with our response and thank for the constructive feedback throughout the review process.

In the revised version of the MS, we have improved the dialogue between supplementary material and Extended Data Figures and the main text: we now link to the Supplementary Material in a more targeted manner and use references to a lower level of index number. All supplementary material chapters and all Extended Data Figures are now called in the main text of the paper. We have also expanded the explanation of the procedure of parameterization in the main text so that this is not missed by readers.

Referee #3 (Remarks to the Author):

I appreciate authors' fair response to my concern and, a such, glad to support publication of this ms.

We thank the referee for the positive evaluation of our work and the feedback in previous review rounds.

Referee #4 (Remarks to the Author):

I thank the authors for addressing the points raised in my review. Regretfully, my reservations remain nearly in full.

We thank the referee for the attention paid to our MS. We address the remaining concerns and hope that the clarifications we offer help resolve the concerns of the referee.

About the possible role of diffusion in the formation of the Dpp gradient

The paper by JPV and colleagues (Stapornwongkul et al., 2020) shows that a functional gradient can be engineered by a synthetic ligand. This demonstrates that diffusion of Dpp may in principle be sufficient to produce a patterned wing. Obviously, a different mechanism may operate in nature and it is important and necessary to explore what is happening in vivo for endogenous Dpp. In this sense, the authors are fully correct to emphasize that this finding does not contradict their own conclusions. Having said this, the authors are not taking seriously enough the key finding from the JPV paper.

1. in their rebuttal, the authors insist that the synthetic signaling ligand engineered by Stapornwongkul et al. (2020) 'is not connected to the formation of the ligand gradient' and that 'with respect to our work, the JPV paper has the value of uncovering a potential role for leakage'. In my view, the Stapornwongkul et al., 2020 provide compelling evidence that diffusion may be sufficient to form a signaling gradient, raising in turn the possibility that simple diffusion of Dpp may likewise be sufficient. Are there clear data excluding this possibility? As explained in my earlier review, the claim by the authors that their modeling approach combined with experiments is totally unbiased, and therefore includes a model whereby diffusion only would operate, is not entirely convincing.

We do not exclude the possibility that a gradient can be made by a mechanism governed by extracellular diffusion. As stated before, our theoretical framework is very broad and this allows it to capture a wide variety of transport models. Moreover, we show that in one of the conditions described in our manuscript (*pentagone* mutants), diffusion plays the central role: the unbound module contributes to 97% of λ^2 and the other transport modules, including the recycling module, have a negligible contribution (Fig. 3b): *pentagone* mutants operate in an ExD-type of regime, where the gradient is formed mainly by extracellular diffusion. As such, our results show that 1) diffusion is indeed sufficient to

form a signaling gradient in vivo (as seen in *pentagone* mutants) and 2) our modeling approach combined with experiments can capture a model whereby diffusion only would operate.
Nevertheless, we find that in large wildtype discs the Dpp gradient is formed by an alternative mechanism, which requires both recycling and extracellular diffusion. Beyond the parameterization work, there is "...clear data excluding [that there is only diffusion]":

i) The effect of downregulation of recycling on the Dpp gradient. In order to downregulate recycling, we have silenced the expression of the recycling Rab proteins: Rab11 and Rab4. We performed RNAi experiments in which we express validated dsRNAs for Rab11 or Rab4 in the posterior compartment of the wing disc. We have measured the recycling rate of Dpp by the nanobody assay upon silencing this Rab proteins and in these conditions the recycling rate of Dpp is significantly decreased. Then, we have observed a decrease of the decay length of eGFP-DppLOP gradient (see Extended Data Figure 6k-m), indicating that recycling plays an important role in gradient formation. Furthermore, when Rab4 is silenced, the eGFP-Dpp^{LOP} gradient does not scale: a mutant in endocytic recycling is a scaling mutant.

ii) The photoconversion and iFRAP assays addressed whether there is recycling of internalized molecules (it could have been that while the parameterization yields a significant recycling, in these two experiments recycling was not detected).

iii) The Dpp^{Timer} assay addresses whether the age of spreading molecules in the extracellular space is compatible with slow phenomena such as intracellular trafficking and recycling (it could have been that the extracellular molecules were younger which would be incompatible with the existence of recycling)

iv) The prerequisite of the free extracellular diffusion mechanism for gradient formation in the work by JPV and colleagues is an important level of leakage in the system. We have investigated experimentally and theoretically the potential role of leakage in our system using a GFP construct which is transcribed, translated, cleaved and trafficked like Dpp. This study suggests that leakage is not an important phenomenon in our system

All these experiments show that recycling is important for the formation of Dpp gradient in large wing.

2. in the main text, the JPV paper is not cited for its main finding, i.e. a functional gradient can be engineered by a synthetic ligand, but rather to emphasize the importance of Dally in gradient formation (p.8, l. 296-8), which is quite misleading. Indeed, the JPV paper does not re-investigate the role of Dally in gradient formation but rather shows that GPI-anchored non-signaling binders, such as Dally for Dpp, contribute to the formation of the signaling gradient of the artificial ligand. Thus, the authors mis-cite this important paper in the field only to strengthen a relatively minor conclusion of their study. This is really not appropriate.

Referee #4 points out that we do not discuss the paper of JPV in sufficient detail for its main findings, namely that:

a) a gradient can be engineered by a synthetic ligand
b) diffusion may be sufficient to form a gradient
c) there might be leakage in the system

In the previous version of our MS, we have considered the article by JPV and colleagues in the main text both in the context of

- a synthetic ligand makes (a) and gradient and there is leakage (c):
  line 76-77: "*Unbound ligand molecules could in principle leak out of the epithelium,* ***as recently studied in a synthetic system where secreted GFP binds to a membrane nanobody and forms gradients***[1]"(added after 1st round of reviews)

- and diffusion may be sufficient to form a gradient (b):
  The combined transport regime that we found in the wing is also different to other regimes described in the literature[1,2]*, that are dominated by the extracellular diffusion of the unbound pool (captured by the unbound module, Extended Data Fig. 1k). In the regimes where the diffusion of the unbound pool dominates transport, morphogen concentration decays as a gradient by means of morphogen capture through irreversible receptor binding*[2] *and/or leakage of the morphogen out of the epithelium, as recently reported*[1]*.* l. 182-188 (added after 2nd round of reviews)

In the revised version of the manuscript, we start the main text (first paragraph) by introducing the controversy on the mechanism of gradient formation (with reference, among others, to the paper of JPV) :

> first paragraph *page 1. To understand how the Dpp gradient decay length is regulated during scaling, it is essential to determine which cellular transport phenomena are important during Dpp gradient formation and which of these change as the tissue grows. Those phenomena could include the extracellular diffusivity of Dpp and the recycling of internalized morphogen molecules,* ***however the relative contribution of these has been the subject of debate[1,3-5] (for a discussion of this debate****, see Supplementary Informtation (SI) 4.2).*

We add another paragraph describing the main findings of the article by JPV and colleagues when discussing free extracellular diffusion-based regimes of morphogen transport:

> *page 6. In the regimes where the diffusion of the unbound pool dominates transport, morphogen concentration decays as a gradient by means of morphogen capture through irreversible receptor binding[2] and/or leakage of the morphogen out of the* epithelium[1]. Indeed, to explore minimal requirements for gradient formation, recent work showed that GFP by itself can form a diffusion-based gradient when binding to membrane associated anti-GFP antibodies, which also limits leakage[1].

We hope that the referee is now satisfied with the level of discussion of the article by JPV and colleagues in our work.

Indirect evidence (# 1)

I maintain that that the binding rates of Dpp to its internalization receptor have not been studied and note that the authors did not address this point in their rebuttal. The general discussion on measurements in science is interesting but slightly off-topic.

Referee #4 comments on the fact that binding rates of Dpp were measured indirectly. What makes powerful our procedure to estimate parameter rates (including the binding rates) is the range of multiple independent assays we have employed. Importantly this is combined with strong statistical approaches to fit the data that provides a precise understanding of what is the actual level of uncertainty in the parameter estimation. Because the model framework we employ is very broad, the conclusions that arrive from our model are solid: Dpp recycles and this contributes to gradient formation. It seems that other reviewers that have assessed this manuscript agree that the parameterization approach we introduce in this work is of high calibre.

Nevertheless, we now explicitly mention that binding rates are measured indirectly in the main text:

> *page 5. Figures 3c,d shows the cluster of value sets determined using this procedure projected on different parameter planes, for the pairs* (k_{on}, k_{off})*,* (k_{on}, D_0) *and* (k, k_{off}) *(see also Extended Data Fig. 4l). It is important to disclaim that some parameters are inferred by our Bayesian approach and not obtained through direct measurement e.g.* k_{on}, k_{off}*. However, given that our model framework is broad and the parameterization supported by five independent quantitative assays, our approach provides parameter value estimates with a narrow level of uncertainty, within an order of magnitude (Fig. 3a,c,d; Extended Data Fig. 4l,m; Extended Data Tables 1,2).*

Otherwise, the authors now provide some evidence that Dally/Dlp modulate the endocytosis of Dpp. This observation is somewhat expected since these HSPGs are known to stabilize Dpp at the cell surface. So, while these data are consistent with the idea that Dally/Dlp act as Dpp internalization receptors, they fail to show that the binding rate of Dpp to its proposed internalization receptor varies in a Pentagon-dependent manner during scaling (abstract).

In the previous round of revision, one of the points of concern raised by this reviewer was the fact that the internalization receptor is unknown and that binding rates of Dpp to its internalization receptor are measured indirectly. We are happy that the referee now agrees that we have addressed the issue of the internalization receptor and seems convinced by our response (for indirect measurements, see comment above “Indirect evidence (#1)”).

However, this reviewer points out now that

« they fail to show that the binding rate of Dpp to its proposed internalization receptor varies in a Pentagon-dependent manner during scaling (abstract) »

Our parameterization of the Pentagone mutant shows indeed that the binding/unbinding of Dpp to its internalization receptor is strongly affected, compatible with the idea that Pentagone regulates the binding/unbinding of Dpp to its internalization receptor. Then, parameterization of control discs of different sizes suggests that this binding/unbinding rates varies during scaling. We also show that, in Dally mutants, the binding/unbinding of Dpp to its internalization receptor is also strongly affected, consistent with the idea that Dally/Dlp could be among the internalization receptors of Dpp. Indeed, it is well possible that there are other internalization receptors of Dpp.

In the sentence of the abstract mentioned by the reviewer, we do not allude to any proposed internalization receptor of Dpp; in particular, not to Dally:

page 1 (abstract). To regulate the contribution of endocytosed Dpp to the spreading extracellular pool during tissue growth, it is the Dpp binding rates that are progressively modulated by the extracellular factor Pentagone, driving scaling.

Moreover, in the main text of the previous version of the MS, we speculated in a very nuanced tone that Dally may be an internalization receptor of Dpp and discuss the potential role of Pentagone on changing the binding properties of Dpp to Dally:

*page 9. We **speculate** that, **if** Dally does function as **one** of the receptors which internalize Dpp, then the k_{on} and k_{off} rates that are regulated by Pentagone **includes** the binding/unbinding rates of Dpp to Dally*

In the new version of the manuscript we softened the tone further:

page 9. We speculate that, if Dally does function as one of the receptors which internalize Dpp, then the k_{on} and k_{off} rates that are regulated by Pentagone could include the binding/unbinding rates of Dpp to Dally, although it may also regulate other unknown receptors instead of Dally

Role of Dpp recycling in gradient formation (# 2)

In their rebuttal, the authors are hammering that they have discovered that 'recycling (of Dpp) is essential for gradient formation' and that 'the contribution of recycling as the disc grows becomes increasingly important'. As mentioned earlier, I respectfully disagree. Introducing the concepts of Modules does not address the extent to which internalized Dpp contribute to gradient formation. Otherwise, the Rab11/Rab4 data are interesting but they do not show that it is the recycling of Dpp that is important. In a diffusion model, gradient formation was shown to depend on the levels of signaling and non-signaling binders (Stapornwongkul et al., 2020). Thus, an alternative interpretation of the Rab11/Rab4 data is that trafficking defects of Tkv and/or Dally/Dlp are causing the Rab11/Rab4 mutant phenotypes.

To summarize, using conditions of downregulated Rab11/Rab4 by RNAi, we have made three key observations:
a) we directly measure the recycling rate of Dpp using the nanobody uptake assay and show that in conditions of Rab11/Rab4 downregulation the recycling rate of Dpp is significantly decreased. It is not just recycling; it is recycling of Dpp;
b) we measure the decay length of GFP-Dpp and show that Rab11/Rab4 downregulation affects Dpp gradient formation;
c) we measure the decay length of GFP-Dpp in discs of different sizes and show that, in discs with downregulated Rab11/Rab4, the Dpp gradient does not scale. These falsification experiments are consistent with the conclusions stemming from the module analysis that recycling is important for gradient formation and gradient scaling.

Referee #4 proposes that an alternative interpretation of the Rab11/Rab4 data is that trafficking defects of Tkv and/or Dally/Dlp are causing the Rab11/Rab4 mutant phenotypes. In the revised version of the manuscript, we do clarify a potential issue of the Rab4/Rab11 RNAi experiment: indeed, alongside its effect on recycling of Dpp, Rab4/Rab11 downregulation could affect trafficking of other components of the pathway such as Dally, Pentagone, Thickveins or others. In any case, whether there is an effect or not on these other factors, our data show that Dpp recycling is itself

affected. If those other factors are affected, they could indeed mediate the effects of Rab11/Rab4 downregulation on Dpp recycling. Indeed, such indirect scenario would imply that the recycling of Dpp ultimately occurs through a mechanism which involves Dally, Pentagone or Thickveins.
We discuss this alternative interpretation of our data in the main text of the revised version of our manuscript:

page 7. Note, that Rab4/Rab11 downregulation could affect trafficking of other membrane components of the pathway (including Dpp receptors, HSPG factors, etc). The effect of this downregulation on recycling of Dpp in this case might be indirectly explained through the trafficking defect of these components.

Model (#3)

I find revealing that the authors highlight the characterization of the transport modules as key novel information on the system. My point remains in full.

Referee #4 finds our conclusion are consequences of our modeling assumptions. Our framework is very general and broad and therefore allows one to capture a very wide variety of transport models. However, it is true, that it does not capture some of transport scenarios previously proposed such as transport via long or monodirectional cytonemes.
In the revised version of the manuscript, we are now mentioning explicitly the limitations of our framework in the main text:

page 3. It is important to be aware of the fact that certain scenarios of transport that have been proposed, such as transport via long or monodirectional cytonemes[6-8], are not captured in our description. Furthermore, unbound ligand molecules could in principle leak out of the epithelium, as recently studied in a synthetic system where secreted GFP binds to a membrane nanobody and forms gradients[1].

Precisions of the measurements (#4)

I am not sure that the quality of the experimental data should be evaluated by their usefulness in defining a parameter range to be used in a specific model.

As the referee has indeed stated, the precision of the estimation of parameters is highly dependent on the reproducibility and precision of the measurements using the transport assays. We therefore took special care in confirming that our measurements are precise: the precision and reproducibility of our transport assays are evaluated by standard errors of the mean and sufficiently high sample sizes (see Figure 2f,g and Extended Data Fig. 3e,g,k; 4b). For FRAP assays, we confirmed that sample sizes in our data set are sufficient for accurate parameterization of transport parameters by studying how the effective diffusivity (estimated by fit from FRAP curve) and goodness of the fit (R^2) stabilizes as we increase the number of individual recovery curves considered for the average FRAP trace (for details, see SI 2.5.1 and Extended Data Fig. 4c,d).
Moreover, we perform an analysis (the robustness analysis SI 3.7) to evaluate how the changes of these measured parameters, within their ranges of uncertainty and even above, affect our conclusions (which regime of transport corresponds to the given set of parameters). This robustness analysis shows that the precision of our measurements is largely sufficient to be confident about the conclusions we draw.

From the comment of this referee, it seems we have failed to clarify the point on the precision of parameter estimation. In order to clarify this point further, we have strengthened the description of chapter 2.5.2 describing the Bayesian computation approach and made all codes available. We have also strengthened the description of the parameter determination procedure and clarified the issue of precision of measurements in the main text of revised version of the manuscript:

page 4. We therefore perform a sequence of data analysis steps using the assays, where each assay adds further constraints to the parameter values.

Briefly, in this sequence of steps, we first fit the dynamics of the nanobody assay to our theoretical framework (SI 2.2.1). This yields narrow confidence intervals for the recycling rate k_r, an effective internalization rate $k_N = k(\frac{k_{on}}{k_{on}+k_{off}})$ and the effective degradation rate k_o (SI 2.2.3). Second, we use Equation 1 and the measured confidence interval of the decay length to constrain the values of D_0, k_{on}, k_{off}. Third, we used an Approximate Bayesian Computation approach to infer posterior distributions for D_0, k_{on}, k_{off}, k_i, k_1 . For this procedure, we sample large sets (on the order of 10^7sets) of parameter values for which we numerically solve the equations that describe Dpp transport under the FRAP assay conditions (SI 1.1), and quantify the similarity between calculated recoveries and the experimental FRAP recovery profiles (SI 2.5.2). This provides narrow ranges for the values of D_0, k_{on}, k_{off}, k_i, k_1; (Extended Data Fig. 4j). Fourth, we further restrict the ranges for the parameters D_0, k_{on}, k_{off}, k_i and k_1 by i) excluding parameter sets that are inconsistent with the measured extracellular fraction (SI 2.5.2; Extended Data Fig. 4k-m) and ii) using the measured value of k_2 obtained from the long-term FRAP assay (SI 2.4; ; Extended Data Fig. 2k-n). The numerical C++ code and the Wolfram Mathematica script for this procedure are available online.

Using this procedure, we obtained a compact cluster of value sets for the eight parameters (see Fig. 3a (orange) for an example in large discs with posterior compartment length $\ell = 144 \pm 4\ \mu m$; Extended Data Fig. 4k). The spread of this cluster corresponds to the precision of the different parameter values. This spread is constrained by the fact that each parameter value set (i.e. each point in the cluster) is consistent with all five assays. It is important to be aware of the fact that the precision of parameter estimation is therefore highly dependent on the quality of measurements performed with each assay (see Figures 2f,g and Extended Data Fig. 3e,g,k; 4b-d and SI 2.5.1 for evaluation of precision of our assays). To address this, we studied how robust the position and spread of the cluster are when we consider a higher imprecision than the one obtained in key assays such as the extracellular fraction assay or the FRAP assay. To this end, we simulated a relaxation in the precision of the assays in our numerical analysis (Extended Data Fig. 5a,b; SI 3.7)). We found that the position and spread of the clusters remained largely unchanged when we consider values beyond the uncertainty range characteristic of these experimental assays. Despite the experimental imprecision (and even when relaxing it further), our parameterization approach yields robust parameter estimates.

Pent (# 5)

the authors have addressed my specific concern. A new PentGFP line (using a MIMIC) was generated and used to show that PentGFP levels increase over time (Fig S15a,b).

We are happy we have clarified this concern of the referee.

Model (# 6)

I understand from the rebuttal that the production rate corresponds to the secretion of 'not previously endocytosed' Dpp by source cells, as estimated from FRAP recovery profiles in Wartlick et al. (2011). However, going back to the original paper, one reads that 'in FRAP, the "production rate" is actually a combination of transcription, translation and secretion, diffusion and degradation in the source, i.e. all the processes that contribute to the flux of molecules out of the source' (pp. 40-41 of the SM in Wartlick et al., 2011). Taking the original definition, my question remains: assuming that Dpp can be internalized by source cells, how does production by source cells of Dpp diffusing into the tissue remains constant if recycling increases over time? If the authors instead assume that Dpp is not internalized by source cells, this assumption needs to be explained. In sum, my point remains.

We thank the referee for the thorough reading of our previous work and for raising this concern. However, let us clarify first bv saying that, in Wartlick et al (2011), the production rate is measured not using FRAP, but by quantifying Dpp transcription over time using BlkRFP. As such, those measurements provide evidence that the *de novo* production of Dpp at the source remains constant over time.

Page 39 of Supplementary Material in Wartlick et al.

To estimate the source width and Dpp production rate during development, we used nuclear GFP driven by the Dpp promoter blk (blk-nGFP, described in SOM EP1). *Since the blk promoter is the Dpp promoter, the GFP fluorescence intensity in the source is then a direct reflection of the Dpp production (transcription) rate $\nu Dpp = \nu nGFP$.*

Assuming that GFP production is adiabatic, i.e. changes in this rate are sufficiently slow, the GFP fluorescence intensity in the source is in a quasi steady state:

$GFP(t) = \nu_{Dpp(t)}/k_{GFP}$ (S-22)

assuming that the GFP degradation rate is constant during development. The Dpp production rate is therefore directly proportional to the GFP fluorescence intensity. ***For further analysis, we therefore used the directly measured GFP fluorescence intensity.***

It is worth noting that in the analysis and fits of the FRAP data in the current manuscript, we have a more detailed theoretical description compared to Kicheva et al. 2007. Now we explicitly define a secretion rate (Eq. S1.1, SI 1.1) that is nonzero only in the source. Recycling in our framework is possible in the entire tissue (source and target). So Dpp can be externalized in source cells due to a constant production rate or due to recycling and receptor unbinding. The contributing flux due to recycling does indeed change over time (both in the source and the target tissue). The previous attempts to fit FRAP mentioned by the reviewer here only consider a single equation and are based on a continuum (no cells are explicitly considered). In this case, the effective secretion rate *per* cell can be computed by estimating a net flux to the target tissue from the source (Kicheva et al). This is not the method we employed here (and not the method used previously to quantify secretion over time in Wartlick et al). Here, we consider the source cells explicitly. Source cells produce Dpp at a constant rate and they also capture, internalize and recycle Dpp the same way target cells do. As such, an effective flux from the source to the target emerges that is indeed a function of the (constant) secretion rate, recycling, diffusion and degradation in the source. Importantly, this flux is an emerging quantity in our current description and follows from the processes at the molecular level that our model captures.

In summary, we consider that Dpp can be internalized and recycled by source cells (Eq. S1.1, SI 1.1) as well as secreted *de novo* at a constant rate. The secretion rate was shown to be constant using assays that quantified transcription of Dpp at the source. As such, there is no contradiction between our previous and present work. We hope that this helps clarify the point raised by the reviewer. In the revised version of the MS, we have now added a clear definition of the production rate in SI 1.1.

Leakage (minor specific point)

The authors have now produced a new secreted GFP, referred to as sGFP(Dpp). In FigS13e, the authors expressed this molecule under the control of a dpp-Gal4 and showed that it accumulates apically far from its source, suggestive of both free diffusion and absence of leakage. These data are nice. Based on these, the authors propose that endogenous Dpp likewise does not leak. This is a surprising conclusion. As pointed out by referee #1 in her/his review, whether Dpp has the same characteristics as extracellular GFP remains an assumption. Again, I am not sure that one can conclude on the basis of these observations that 'the Dpp morphogen system leakage is very small if at all existing' (rebuttal).

Referee #4 is not convinced that our experiments using GFP processed and secreted like Dpp address the question on Dpp leakage itself.
The point we make with this new reagent which is transcribed, translated, furin-cleaved and secreted like Dpp, is that a molecule which is processed like Dpp is in an extracellular compartment from where leakage is negligible. In the main text of the revised version of the manuscript, we now precise that the results on leakage we obtain using GFP, concern molecules processed and secreted in the epithelium like Dpp (and not directly Dpp itself):

page 3. This suggests that for molecules processed and secreted in the epithelium like Dpp, leakage is unlikely to be important (Extended Data Fig. 2a-j; see further considerations in SI 2.11, SI 3.10, SI 4.3).

Otherwise, the observation that sGFP(Dpp) does not form a gradient is perhaps expected and not particularly meaningful given the absence of specific receptors for this artificial ligand (see Stapornwongkul et al., 2020).

This experiment was designed specifically to address whether there is leakage. Therefore it is important to perform this experiment in the absence of receptors that could impair leakage by themselves (see Stapornwongkul et al., 2020). This experiment indeed tells us that, for a molecule processed like Dpp, leakage is not important.

Minor comment: the results shown in S13e would need to be better described: is this a fixed sample? What are the cells stained with phalloidin that are locate apically? If it is a cross-section along the a-p axis, why is the expression of sGFP(Dpp) discontinuous at the source? Also, the distribution of sGFP appear heterogeneous on the surface view (right panels), with both cell-to-cell variability and specific accumulation at cell-cell junctions: please comment.

We thank the referee for pointing out that the description of the Extended Data Figure 2e is insufficient. In this figure, the disc is fixed and stained with F-actin phalloidin to evaluate the apicobasal localization of $sGFP^{Dpp}$. As confirmed by apical colocalization with phalloidin, most $sGFP^{Dpp}$ is seen in a thin region within the epithelium in the apical side. Above the apical side of columnar cells of the wing disc, are found the cells of the peripodial membrane: these are the cells stained with phalloidin that are located apically. Occasionally, at the edge of the source, Dpp-gal4 drives expression showing cell-to-cell variability. This explains the discontinuity of $sGFP^{Dpp}$ expression that can be seen in the orthogonal view at the edge of the source.
Right panels of Extended Data Figure 2e present an xy single plane of the sample. On the xy plane of the $sGFP^{Dpp}$ image, the localization of $sGFP^{Dpp}$ at cell-cell junctions can be observed, confirming that $sGFP^{Dpp}$ is localized in the extracellular space of the epithelium. The cell-to-cell variability of sfGFP fluorescence can therefore be explained by the fact that most of $sGFP^{Dpp}$ can be seen in a thin (1 µm thick) region within the epithelium in the apical side and the maximum of $sGFP^{Dpp}$ fluorescence intensity in different cells is outside of the x-y plane chosen in the figure.
In the revised version, we have now strengthened the description on Extended Data Figure 2e in the corresponding figure legend.

Last, the authors 'focused on the target cells in the posterior compartment of the wing imaginal disc' (l. 58). A very recent paper from the Affolter lab (bioRxiv preprint doi: https://doi.org/10.1101/2020.11.23.394379) provides a rationale for this otherwise arbitrary choice. It would be useful for non-expert readers to cite this work.

We thank the reviewer for pointing out that our choice of focusing on the target cells has not been rationalized for non-expert readers. We focus on cells in the posterior compartment, as cells in this compartment are target cells-only and do not mix with source cells. In the revised version of the manuscript we precise this and cite the corresponding references.

page 2. We focused on cells in the posterior compartment of the wing imaginal disc, as this compartment is exclusively composed of target cells which do not produce Dpp themselves[9-11].

Referee #5 (Remarks to the Author):

Review for “Morphogen gradient scaling by recycling of intracellular Dpp”

To give a quick context, I am writing this as a “fresh” Referee after the previous rounds of review and revision. I have read thoroughly the latest draft of the manuscript and the multiple rounds of referee reports and author rebuttals. After a brief introduction, I give my Referee report on the manuscript and also the rebuttals to points raised by referees.

We thank the referee for the time to assess our MS and past reviewer comments.

The question of morphogen scaling has been a long-standing one and has been challenging to find clear mechanisms. Even in the apparently simple morphogen Bicoid, there is significant debate over scaling – including if it even happens. Therefore, the aims of this paper are laudable and likely to have a broad impact.

Overall, the central results of the paper are highly impressive. The range of quantitative experiments, combined with theory, are very powerful and help to provide evidence for a plausible mechanism of morphogen scaling. The genetics and controls generally seem performed to a high standard, giving confidence in the main conclusions.

Below, I discuss a number of the issues that have arisen during the review process and outline my thoughts regarding the author responses. I also raise a few further issues that should be addressed.

We are pleased that the referee appreciates the quality and impact of our work.

Major comments:

1. There is significant debate between the authors and the referees (particularly #2 and #4) about the modelling and what can be interpreted from this. I break this down below:

a. “Indirect measurements”: Referees raise the concern that parameters are indirectly inferred and heavily reliant on the model choice. Relatedly, this then asks questions about the falsifiability of the model.
This is a valid concern, and it is certainly true that too many studies are not rigorous enough in how they incorporate model fitting and then how such parameter values are interpreted. A particular challenge in biology is that the range of mechanisms are so varied that it can be difficult to provide a definitive model.

Having said this, I am generally convinced by the Authors responses (with some important caveats below). The range of multiple independent assays to fit parameters is very powerful. Their statistical approach appears valid and provides a suitable level of uncertainty. Of course, their conclusions are model dependent, but they consider a broad model framework (though see (1.b)). Therefore, I am happy that the paper conclusions – e.g. the level of recycling and the role this plays in gradient scaling – are on a solid footing.

We are pleased that the referee appreciates our work and appreciate the detailed assessment and feedback.

b. “The generality of the model”: There is substantial debate through the previous rounds of revision about whether the model was being suitably general enough. The authors in the main paper state: “This description is equanimous: depending on parameter values, the same general description allows to capture a wide variety of transport models.” Overall, with the below caveat, this is a fair statement and the results as presented in the main paper are generally convincing with regards to this.

We are happy that the referee agrees that our approach is "equanimous" and address the caveat the referee mentions below.

However, in the response to reviewers, the Authors are somewhat bolder, e.g. response to Referee 2: "Therefore, our conclusions are not a consequence of particular modeling assumptions but are uniquely obtained in a single phenomenological framework". This is not entirely correct. As an example, as highlighted by Referee 4, the authors do not consider cytoneme transport (Roy et al Science 2014). I believe in their original response they claim that cytoneme transport can be subsumed within the current model: e.g. "extracellular diffusion incorporates not just the Brownian motion of molecules, but also their binding and retention by the extracellular matrix". This is insufficient for cytoneme mediated transport, which cannot be described (at least in a simple manner) within the authors existing framework. For an example of modelling cytoneme transport modelling see: Improving the understanding of cytoneme-mediated morphogen gradients by in silico modeling, PLoS Computational Biology 2021.

Of course, cytoneme transport of Dpp in the wing disc remains contentious. Yet, it definitely merits mentioning as a possible oversight in the paper. I note that, to the best of my knowledge, the paper does not cite any cytoneme-mediated transport papers.
A related factor here is one of dimensionality. The Authors present a 1D spatial model. Yet, once one starts considering cytonemes (or other possible directed transport mechanisms) it is important to distinguish between the dimensionality of the transport mechanism; e.g. extracellular diffusion is effectively 3D whereas cytoneme transport is effectively 1D. Such considerations are currently not considered in the model. I don't expect new analysis at this point, but clearer discussion of limitations is merited.

We would like to thank the referee for this comment. We agree that cytoneme-based transport merits to be mentioned in the context of the theoretical framework we have employed and corresponding references to be cited. In the revised version of the manuscript, we now explicitly mention the limitations of our framework in the main text and point the reader to the related literature:

> *page 3. It is important to be aware of the fact that certain scenarios of transport that have been proposed, such as transport via long or monodirectional cytonemes[6-8], are not captured in our description. Furthermore, unbound ligand molecules could in principle leak out of the epithelium, as recently studied in a synthetic system where secreted GFP binds to a membrane nanobody and forms gradients[1].*

In summary, I am happy with the Authors response to the valid critiques raised – particularly by Referee 2 and in both rounds of revision by Referee 4 - about model fitting and how such fitting can be interpreted. Their conclusions are as solid as can be, given current experimental limitations. However, I feel clearer caveats need to be included within the manuscript to make the above issues clearer.

We are pleased that this referee's assessment is that our response to both referee 2 and 4 is convincing, as these were the major remaining concerns raised in the review process. We also appreciate this referee's comment about the strength of our conclusions ("Their conclusions are as solid as can be, given current experimental limitations."). Werespond to the concerns and suggestions of this referee below.

2. Regarding specific issues raised by Referee #2:

a. Discussion of the Fried and Iber, Nat Comm 2014, I am satisfied by the Author response. The range of FRAP experiments are quite convincing with regards to the relative equilibrium state of the gradient.

b. Approximation given by Eq. 1. I am happy with the Authors response.

c. Improved pulse experiments and discussion. This has been adequately addressed (though see my further comments below).

d. The response to the Referee regarding the FRAP experiment analysis is generally convincing. However, the Authors state in their original response: "Dpp transport depends on eight parameters. Three of them (endocytosis, recycling, clearance

and immobilization rate) are determined by means of the nanobody uptake assay (explained in Supplementary Material Section 2.2). A fourth rate, the degradation rate of the immobile pool, was determined by measuring the slow relaxation rate in a long-term FRAP experiment. The last three, the extracellular diffusion and the binding/unbinding rate were determined by means of the Monte Carlo sampling and Approximate Bayesian Computation (Supplementary Material Section 2.5.2).” Unless I’m missing something, this text in fact only refers to seven, not eight, parameters. I suspect this is a typo, but can the Authors clarify.

We apologize for this confusion.

In summary, we first fit the dynamics of the nanobody assay to our theoretical framework (SI 2.2.1). This yield narrow confidence intervals for the recycling rate $\mathrm{k_r}$, an effective internalization rate $\mathrm{k_N} = \mathrm{k}(\frac{\mathrm{k_{on}}}{\mathrm{k_{on}}+\mathrm{k_{off}}})$ and the effective degradation rate $\mathrm{k_o} = \mathrm{k_1} + \mathrm{k_i}$ (SI 2.2.3). Second, we use the measured value of the decay length with Equation 1 to obtain an expression for the diffusion constant $\mathrm{D_0}$ as a function of the remaining unknown parameters. Third, we numerically solve the equations that describe Dpp transport under the FRAP assay conditions (SI 1.1), and fit the theoretical recoveries to the experimental FRAP data using an Approximate Bayesian Computation approach (SI 2.5.2). This provides reduced ranges for the values of ($\mathrm{D_0}$, $\mathrm{k_{on}}$, $\mathrm{k_{off}}$, $\mathrm{k_i}$, $\mathrm{k_1}$; Extended Data Fig. 4j). Fourth, we further restrict the possible values for $\mathrm{D_0}$, $\mathrm{k_{on}}$, $\mathrm{k_{off}}$, $\mathrm{k_i}$, $\mathrm{k_1}$ by excluding parameter sets that are not consistent with the measured extracellular fraction (SI 2.5.2; Extended Data Fig. 4k ; 4l,m) and using the direct estimate for $\mathrm{k_2}$ from the long-term FRAP assay (SI 2.4).

Overall 8 parameters are fitted.

In the description mentioned by the referee above, the “hidden” parameter comes from deconvolving the effective degradation rate ($\mathrm{k_o} = \mathrm{k_1} + \mathrm{k_i}$) into the clearance rate of the mobile pool ($\mathrm{k_1}$) and the immobilization rate ($\mathrm{k_i}$).

In the revised version of the manuscript, we make description of the fitting routine and Approximate Bayesian Computation approach more visible both in the supplementary information (chapter 2.5.2) and the main text:

page 4. We therefore perform a sequence of data analysis steps using the assays, where each assay adds further constraints to the parameter values.

Briefly, in this sequence of steps, we first fit the dynamics of the nanobody assay to our theoretical framework (SI 2.2.1). This yields narrow confidence intervals for the recycling rate k_r, an effective internalization rate $k_N = k(\frac{k_{on}}{k_{on}+k_{off}})$ and the effective degradation rate k_o (SI 2.2.3). Second, we use Equation 1 and the measured confidence interval of the decay length to constrain the values of D_0, k_{on}, k_{off}. Third, we used an Approximate Bayesian Computation approach to infer posterior distributions for D_0, k_{on}, k_{off}, k_i, k_1 . For this procedure, we sample large sets (on the order of 10^7 sets) of parameter values for which we numerically solve the equations that describe Dpp transport under the FRAP assay conditions (SI 1.1), and quantify the similarity between calculated recoveries and the experimental FRAP recovery profiles (SI 2.5.2). This provides narrow ranges for the values of D_0, k_{on}, k_{off}, k_i, k_1; (Extended Data Fig. 4j). Fourth, we further restrict the ranges for the parameters D_0, k_{on}, k_{off}, k_i and k_1 by i) excluding parameter sets that are inconsistent with the measured extracellular fraction (SI 2.5.2; Extended Data Fig. 4k-m) and ii) using the measured value of k_2 obtained from the long-term FRAP assay (SI 2.4; ; Extended Data Fig. 2k-n). The numerical C++ code and the Wolfram Mathematica script for this procedure are available online.

e. I fully agree with the Referee #2 that, given the large-scale extensive analysis, that all appropriate scripts should be made fully available. I hope the Authors generate a Github (or similarly accessible) repository for these codes. It would also be excellent if the authors also include the quantitative data sets. Then, other researchers have the freedom to test different analysis approaches or different models. This will substantially increase the impact of the paper and sets a clear standard for presentation of results in quantitative biology papers.

We have now uploaded the C/C++ and Wolfram Mathematica scripts used for the parameter estimation in Github (https://github.com/zenah12/DppTrafficking-/blob/main/README.md) and will make raw data available for the main figures. Other raw data will also be available upon upon request.

f. I agree that the Authors have satisfactorily dealt with the “small points” raised by the Referee #2.

We are pleased that the reviewer appreciates our response to Reviewer #2.

3. Related to a lot of the above discussion, Referee #4 highlights potential problems with the data interpretation and the subsequent interpretation. In particular: "However, the extent to which internalized Dpp DOES contribute to gradient formation remains unaddressed. Despite this, the authors did not change their conclusion ("therefore recycling contributes to gradient formation"). Thus, the issue remains." I understand the Referee point and think that such scepticism is justified. However, I am overall convinced by the Author response to this issue (given the caveat that I think the limitations in the approach needs to be made clearer in the main paper). I find the range of quantitative experiments genuinely impressive, and it's hard to see what more could be done given current experiment and imaging tools. Of course, future work using more sophisticated approaches may show the conclusions here are not exactly correct, but that's science.

In the revised version of the manuscript, we now explicitly discuss this limitation of our framework in the main text:

page 3. It is important to be aware of the fact that certain scenarios of transport that have been proposed, such as transport via long or monodirectional cytonemes[6-8], are not captured in our description. Furthermore, unbound ligand molecules could in principle leak out of the epithelium, as recently studied in a synthetic system where secreted GFP binds to a membrane nanobody and forms gradients[1].

4. There is significant discussion of the precision of parameter measurement. There are two major points here: (1) the Authors use sensible approaches and provide suitable ranges on parameters; (2) the description in the Methods, as highlighted by Referee #4 are currently insufficient. I did not find the description of sufficient depth (e.g. Supp Mat Part 2, 2.5.2). This would definitely benefit from publishing of the code alongside the paper, as currently the presented information is insufficient for another researcher to independently validate these results. Given that such output is central to the paper conclusions, such information (the code and the quantitative numbers used for fitting) should be made available.

In order to clarify the point on the precision of parameter estimation, this referee therefore proposes to strengthen the description of chapter 2.5.2 describing the Bayesian computation approach and to make all codes available. In the revised version of the manuscript we have addressed both of this two proposals of referee #5.
Moreover, we have now clarified the issue of precision of measurements in the main text of revised version of the manuscript:

page 4. Using this procedure, we obtained a compact cluster of value sets for the eight parameters (see Fig. 3a (orange) for an example in large discs with posterior compartment length $\ell = 144 \pm 4$ μm; Extended Data Fig. 4k). The spread of this cluster corresponds to the precision of the different parameter values. This spread is constrained by the fact that each parameter value set (i.e. each point in the cluster) is consistent with all five assays. It is important to be aware of the fact that the precision of parameter estimation is therefore highly dependent on the quality of measurements performed with each assay (see Figures 2f,g and Extended Data Fig. 3e,g,k; 4b-d and SI 2.5.1 for evaluation of precision of our assays). To address this, we studied how robust the position and spread of the cluster are when we consider a higher imprecision than the one obtained in key assays such as the extracellular fraction assay or the FRAP assay. To this end, we simulated a relaxation in the precision of the assays in our numerical analysis (Extended Data Fig. 5a,b; SI 3.7)). We found that the position and spread of the clusters remained largely unchanged when we consider values beyond the uncertainty range characteristic of these experimental assays. Despite the experimental imprecision (and even when relaxing it further), our parameterization approach yields robust parameter estimates.

We have also made the code written for the different steps described above available, together with a description of how the code can be used to obtain parameter values. The code, data and details of the computational methods can be accessed here: https://github.com/zenah12/DppTrafficking- and the link is made available in the supplementary material together with the description of the fitting routine.

5. Paper structure and format:

a. I found reading the paper very challenging. As highlighted by Referee #1, the paper is extremely dense and hard to parse at times. I appreciate the length limitation imposed by the journal, but I think that the paper itself can still be substantially

improved in terms of readability. I suspect by this stage that the Authors are tired of editing it, but in its current form I don't think the paper will have the impact it deserves due to the very dense writing.

In the revised version of the MS, we have improved the dialogue between supplementary information and Extended Data Figures and the main text: we now link to the Supplementary Material in a more targeted manner and use references to a lower level of index number. All supplementary information chapters and all Extended Data Figures are now called in the main text of the paper. We have also expanded the explanation of the procedure of parameterization in the main text so that this is not missed by readers.

b. Related to the above point, the Supplementary Material Section II is not clear enough. To properly understand this paper, requires a reading of the Supplementary Material. Yet, this material is very dense. Notably, the page numbers of the index do not match the actual page numbers, making navigation through the document more difficult than necessary. As with the paper itself, this material needs to be gone through properly and brought up to standard.

We have worked to improve the accessibility of both the main text and the supplementary information in the revised version of the manuscript (see below).
We have thoroughly double-checked the numbering of the table of contents and found no mismatch. It is possible that the mismatch in the document available to the reviewer was caused by the automatic compilation produced by the website of the journal.

c. The Extended Figures should be ordered sequentially as they appear in the main text. It was frustrating moving back and forth through a complex series of documents.

We have now ordered the Extended Data Figures according to call in main.

In summary on this point, I found, with a lot of hard reading (and having five different documents related to the paper open on my computer at the same time) that the paper was rewarding and has very interesting results. However, in its current form, it is likely to lose some impact due to a degree of impenetrability.

We have worked to improve the accessibility of both the main text and the supplementary information in the revised version of the manuscript. In summary, we have made the following changes to make the work more accessible and improve readability:

i) We now call chapters in supplementary information using the lowest index level to direct the reader straight to the relevant material. In addition, all supplementary information chapters and EDFigs are now called in the main text. We hope that this improves the readability of our work and helps the reader obtain the relevant information.
ii) We now provide a brief, albeit an in depth, explanation of the parameterization method in the main text.

page 4. Briefly, in this sequence of steps, we first fit the dynamics of the nanobody assay to our theoretical framework (SI 2.2.1). This yields narrow confidence intervals for the recycling rate k_r, an effective internalization rate $k_N = k(\frac{k_{on}}{k_{on}+k_{off}})$ and the effective degradation rate k_o (SM 2.2.3). Second, we use Equation 1 and the measured confidence interval of the decay length to constrain the values of D_0, k_{on}, k_{off}. Third, we used an Approximate Bayesian Computation approach to infer posterior distributions for D_0, k_{on}, k_{off}, k_i, k_1 . For this procedure, we sample large sets (on the order of 10^7 sets) of parameter values for which we numerically solve the equations that describe Dpp transport under the FRAP assay conditions (SI 1.1), and quantify the similarity between calculated recoveries and the experimental FRAP recovery profiles (SI 2.5.2). This provides narrow ranges for the values of D_0, k_{on}, k_{off}, k_i, k_1; (Extended Data Fig. 4j). Fourth, we further restrict the ranges for the parameters D_0, k_{on}, k_{off}, k_i and k_1 by i) excluding parameter sets that are inconsistent with the measured extracellular fraction (SI 2.5.2; Extended Data Fig. 4k-m) and ii) using the measured value of k_2 obtained from the long-term FRAP assay (SI 2.4; ; Extended Data Fig. 2k-n). The numerical C++ code and the Wolfram Mathematica script for this procedure are available online.

iii) We have now ordered EDF according to call in main

iv) The code used for the parametrization method is now accessible online (https://github.com/zenah12/DppTrafficking-).
v) Supplementary sections 1 and 2 are now merged into a single document to allow the reader to access information at once.
vi) Index pages are correct and the supplementary information is restructured for better penetrability
vii) We provide further detail in our description of the parametrization method in Section 2.5.2 in the Supplementary material.

6. Tandem reporter results:

a. I was surprised reading the results related to the tandem reporter that there was no mention (that I found) of how background signal was accounted for. The background signal can make a big difference to the measured ratio. In Durrieu et al MSB 2018, which used the tandem reporter to explore Bicoid dynamics, they used linear unmixing to provide background estimates for each individual sample. Here, how was background accounted for in the tandem reporter and what effect does this have on the reported ratio?

We would like to thank the reviewer for spotting this. In fact, we have omitted a mention of the background in this experiment. Indeed, we did not include in the description of the measurement of our Dpp^{timer} assay the procedure of background removal. To determine G_{ext}, R_{ext}, G_{int} and R_{int}, the corresponding background has been subtracted. Background intensity has been measured at the edge of the corresponding gradients ($G_{ext}(x)$, $R_{ext}(x)$, $G_{int}(x)$ and $R_{int}(x)$). We have now corrected this in chapter 2.6.2 SI2.
In general, when measuring Dpp gradients, we considered background in two different ways (as described in chapter 2.1.2): i) by measuring the fluorescence at the edge of the corresponding gradient and ii) by estimating background as a fitted offset parameter in the exponential fit to the gradient.
For the Dpp^{timer} analysis, we have in addition performed linear unmixing analysis. This analysis showed that subtracting the measured background at the edge is equivalent to linear unmixing. This is because of two reasons: i) the background is largely homogeneous in the disc, and ii) the length of the posterior compartment is ten-fold longer (approximately) than the decay length of the gradient (unlike the case of the Bicoid gradient, which is only half of the length of the embryo approximately). Therefore, fluorescence at the edge of the disc is background. We disclaim this in the revised version (chapter 2.1.2):

> ***Background substraction.*** *When measuring Dpp gradients, we considered background in three different ways: i) by measuring the fluorescence at the edge of the corresponding gradient, ii) by estimating background as a fitted offset parameter in the exponential fit to the gradient and (iii) linear unmixing analysis[12]. This analysis showed that substracting the measured background at the edge is equivalent to linear unmixing. This is because of two reasons: i) background is largely homogeneous in the disc, and ii) the length of the posterior compartment is ten-fold longer (approximately) than the decay length of the gradient (unlike the case of the Bicoid gradient, which is only half of the length of the embryo approximately). Therefore, fluorescence at the edge of the disc is background.*

b. This is a minor point, but the bead assay is not necessarily giving the precise ratio between the two fluorophores (though it is likely close approximation). When imaging live, there are additional factors such as scattering through the tissue that affect the GFP and mKate2 channels differently. The ratio value F can be left as an additional fitting parameter. This would increase uncertainty in fitting parameters but I doubt this will make a large change.

We thank the reviewer for pointing this out. Indeed, the ratio F does not capture all the properties of the two fluorophores in the tissue. We have been aware of this issue and are now explicitly disclaiming it in the revised version. Indeed, for this reason, in the analysis of the timer data to study the relative ages of molecules in the intracellular *versus* the extracellular space (Fig. 4o), we have chosen to use the ratios of relative ages to avoid this issue already in the original version. When we use ratios, the calibration parameter F cancels out and the analysis becomes independent of calibration.

We however used F for Extended Data Figures 6e,f. Here, the point was to perform important controls to validate that the sfGFP-mKate2-Dpp molecule does behave like a timer. If sfGFP-mKate2-Dpp is a timer we expect that, i) the spatial profile of sfGFP fluorescence and mKate2 fluorescence to be different, ii) their ratios to be lower in the source than in

the target tissue because in the source there is an additional pool of young molecules in the secretory pathway and iii) their ratios to not be homogenous in the target, but graded, reflecting the older age of molecules as a function to the distance from the source boundary. By itself, this analysis does not in principle need the calibration factor F and could be simply performed using the fluorescence levels of the two fluorophores directly. We have however made an effort to present the data as closely as possible to a "real" concentration of molecules instead of just fluorescence. To this end, we determined a calibration factor F by means of beads and used it for Extended Data Figures 6e,f. We are however aware of the limitations of this factor that does not capture the optical properties of the fluorophores in the tissue.

In the revised version, we have disclaimed the limitations of F, and, to avoid confusion, we simply divide the fluorescence of mKate2 by the calibration factor F (New Extended Data Figure 9e,f):

page 40 of Supplementary Material (Controls of the timer): It is worth disclaiming that F does not capture potential differences in the fluorescence behaviour (e.g. scattering) between sfGFP and mKate2 in the wing tissue and therefore is only an approximation to have an idea of the real concentration of fluorescent molecules in our system.
This way, the adjusted fluorescence intensity profiles for sfGFP ($g^*(x)$) and mKate2 ($r^*(x)$) (Extended Data Fig. 6d), which proportional to the respective concentration profiles can be written as

$$g^*(x) = G(x) \qquad \text{(S2.1)}$$

$$r^*(x) = 1/F \times R(x) \qquad \text{(S2.2)}$$

The ratio between the $r^*(x)$and $g^*(x)$ profiles is indicative of the relative age $A(x)$ of Dpp molecules as a function of position in the gradient. Note that closer to the source, there are less mature mKate2 molecules, confirming that Dpp molecules are younger. The relative amounts of sfGFP and mature mKate2 converge as we move away from the source, consistent with an increase in $A(x)$ with the distance to the source (Extended Data Fig. 6d-f). The value of F does not impact the spatial profile of $A(x)$.

7. Framing of the results: I appreciate that for a journal such as Nature that the results must be seen to have broad implication. It is important to note I think the results presented here are of sufficient interest and breadth to be suitable for Nature. Having said that, there are times in the paper where I feel some "overselling" occurs.

a. The abstract states: "Here we unravel the cell biological basis behind scaling…" . I feel a more accurate statement would be along the lines of "Using formation of the Dpp gradient in Drosophila we unravel a cell biological basis behind scaling." Essentially, at a minimum I ask that "the" in the original sentence be changed to "a".

Agreed. We have edited the abstract accordingly:
page 1. « Here, focusing on the Dpp gradient in the Drosophila wing disc, we unravel a cell biological basis behind scaling.»

b. In the abstract: "… which ultimately allows adaptation of shape and pattern to different sizes of organs in different species." This is too bold and needs to be more balanced. The initial part of the sentence does have "could", but this sentence currently reads as claiming more certainty than is there given the results presented are only for one morphogen in one species.

Agreed. We have edited the abstract accordingly:
page 1. « Thus, for some morphogens, evolution may act on endocytic trafficking to regulate the range of the gradient and its scaling, which could allow adaptation of shape and pattern to different sizes of organs in different species. »

c. First intro paragraph: "To understand how the gradient decay length is regulated during scaling, it is essential to determine which cellular transport phenomena are important during Dpp gradient formation and which of these change as the tissue grows." This sentence jumps from a broad point to a specific case – that of Dpp. In its current wording, it could be misread to mean that by simply understanding Dpp gradient formation we can understand gradient formation of all morphogens, which is clearly incorrect. I suggest splitting this into two distinct sentences – one on the general question of understanding decay length scaling, and a second on using Dpp as a specific example to tackle this problem.

Again, agreed and have provided edits:

page 1. « Morphogen gradient scaling implies the regulation of the gradient range (i.e. its decay length) to remain proportional to tissue size[13,14]. In the case of the Dpp morphogen gradient, to understand how the decay length is regulated during scaling, it is essential to determine which cellular transport phenomena are important during Dpp gradient formation and which of these change as the tissue grows. »

d. "Therefore, recycling contributes to gradient formation and could in principle underlie scaling." This sentence is too bold at this point, given the evidence that has been presented up to this point. I would suggest removing or reworking to be more balanced (nearly the same sentence at the end of the manuscript would be fine).

Our edits:

page 2. « The photoconversion experiment suggests that recycling could contribute to some extent to gradient formation and might in principle underlie scaling. »

e. "In our experiment, the measured Dpp Timer age in the extracellular and intracellular space is the same". Given the experimental uncertainty (large error bar), it is not appropriate to say "is the same"; "are similar" feels more appropriate.

Our edits:
page 7. « In our experiment, the measured Dpp^{Timer} age in the extracellular and intracellular space are similar »

Minor comments

- Line 48 typo (missing "a") - "…such a gradient is formed."
Done

- Paragraph starting line 48. Two sentences starting "Because…" . This reads as very clunky.
Done

- Fig. 1d is referenced prior to Fig. 1c
Done

- I appreciate space is an issue but Fig 2f and Fig. 2g should have y-axis labels even if stated in legend

Done

- "Each of these assays depends on most transport rates. We therefore perform a sequence of steps using the assays, where each assay adds further constraints to the parameter values (SM 2.5.2)." I find this sentence too vague. At least give the central idea of the fitting, as it is essential to the paper conclusions.
Done

- Line 134, should be Fig. 3a,b I think

- At the point that Fig. 3a is introduced, there is no mention of the different measurements in wings of different size. Therefore, when first reading it is very surprising to see "l=144" and "l=80" as it's not clear what this means. It would help to have a sentence before this stating that measurements were done at two general wing sizes.
Done

- "Pentagone mediates this shift." This is a somewhat odd sentence – no references or data are presented to back up this statement.
page 8. Our analysis below suggests that Pentagone mediates this shift.

- "We show here that scaling of Dpp gradients in dally mutant discs perturbed", should be "We show here that scaling of Dpp gradients in dally mutant discs is perturbed"

Done

- Figure 2 legend, "ℓ = 144.6 ± 4 µm" – the level of significance on the two numbers is not the same. I assume it should be ±4.0.

Done

In conclusion, this is, as the Authors themselves state, a "tour de force" to tackle the mechanism for how Dpp scales. This provides important new insight into an essential process in development. It will be suitable for publication with some careful rewriting and cleaning up the last remaining issues highlighted above and by the other Referees.

We are very pleased to have received this independent validation of the quality and impact of our work. We hope that the careful clarifications, additional information and adjustments in our tone have addressed the referee's concerns.

1 Stapornwongkul, K. S., de Gennes, M., Cocconi, L., Salbreux, G. & Vincent, J. P. Patterning and growth control in vivo by an engineered GFP gradient. *Science* **370**, 321-327, doi:10.1126/science.abb8205 (2020).

2 Zhou, S. *et al.* Free extracellular diffusion creates the Dpp morphogen gradient of the Drosophila wing disc. *Curr Biol* **22**, 668-675, doi:10.1016/j.cub.2012.02.065 (2012).

3 Entchev, E. V., Schwabedissen, A. & Gonzalez-Gaitan, M. Gradient formation of the TGF-beta homolog Dpp. *Cell* **103**, 981-991, doi:10.1016/s0092-8674(00)00200-2 (2000).

4 Kicheva, A. *et al.* Kinetics of morphogen gradient formation. *Science* **315**, 521-525, doi:10.1126/science.1135774 (2007).

5 Zhu, Y., Qiu, Y., Chen, W., Nie, Q. & Lander, A. D. Scaling a Dpp Morphogen Gradient through Feedback Control of Receptors and Co-receptors. *Dev Cell* **53**, 724-739 e714, doi:10.1016/j.devcel.2020.05.029 (2020).

6 Aguirre-Tamaral, A. & Guerrero, I. Improving the understanding of cytoneme-mediated morphogen gradients by in silico modeling. *PLoS Comput Biol* **17**, e1009245, doi:10.1371/journal.pcbi.1009245 (2021).

7 Rojas-Rios, P., Guerrero, I. & Gonzalez-Reyes, A. Cytoneme-mediated delivery of hedgehog regulates the expression of bone morphogenetic proteins to maintain germline stem cells in Drosophila. *PLoS Biol* **10**, e1001298, doi:10.1371/journal.pbio.1001298 (2012).

8 Roy, S., Huang, H., Liu, S. & Kornberg, T. B. Cytoneme-mediated contact-dependent transport of the Drosophila decapentaplegic signaling protein. *Science* **343**, 1244624, doi:10.1126/science.1244624 (2014).

9 Akiyama, T. & Gibson, M. C. Decapentaplegic and growth control in the developing Drosophila wing. *Nature* **527**, 375-378, doi:10.1038/nature15730 (2015).

10 Matsuda, S. *et al.* Asymmetric requirement of Dpp/BMP morphogen dispersal in the Drosophila wing disc. *bioRxiv*, 2020.2011.2023.394379, doi:10.1101/2020.11.23.394379 (2021).

11 Posakony, L. G., Raftery, L. A. & Gelbart, W. M. Wing formation in Drosophila melanogaster requires decapentaplegic gene function along the anterior-posterior compartment boundary. *Mechanisms of Development* **33**, 69-82, doi:https://doi.org/10.1016/0925-4773(90)90136-A (1990).

12 Durrieu, L. *et al.* Bicoid gradient formation mechanism and dynamics revealed by protein lifetime analysis. *Mol Syst Biol* **14**, e8355, doi:10.15252/msb.20188355 (2018).

13 Inomata, H. Scaling of pattern formations and morphogen gradients. *Dev Growth Differ* **59**, 41-51, doi:10.1111/dgd.12337 (2017).

14 Romanova-Michaelides, M., Aguilar-Hidalgo, D., Julicher, F. & Gonzalez-Gaitan, M. The wing and the eye: a parsimonious theory for scaling and growth control? *Wiley Interdiscip Rev Dev Biol* **4**, 591-608, doi:10.1002/wdev.195 (2015).

Reviewer Reports on the Third Revision:

Referee #5 (Remarks to the Author):

I appreciate the efforts that the authors have undertaken to satisfy the latest round of reviews. I do not see any major issues with the current version. As already discussed, the paper itself is very dense and can be challenging at times to parse, but that's partly due to the large depth of work presented. The caveats are a big improvement and now the manuscript presents a more balanced view.

Timothy Saunders