UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
CURSO DE GRADUAÇÃO EM BIOMEDICINA
TRABALHO DE CONCLUSÃO DE CURSO

Redimensionamento de camadas germinativas embrionárias durante a evolução de *Drosophila* - Investigação da dinâmica do gradiente de Dorsal

PRISCILLA AMBROSI

Orientadoras:
Profa. Dra. Vera Lúcia da Silva Valente Gaiesky (UFRGS)
Profa. Dra. Cláudia Mieko Mizutani (Case Western Reserve University)

Co-orientador:
Prof. Dr. Hillel Chiel (Case Western Reserve University)

Porto Alegre
Julho/2013

PRISCILLA AMBROSI

Redimensionamento de camadas germinativas embrionárias durante a evolução de *Drosophila* - Investigação da dinâmica do gradiente de Dorsal

Trabalho de conclusão de curso de graduação apresentado ao Instituto de Ciências Básicas da Saúde da Universidade Federal do Rio Grande do Sul, como requisito parcial para obtenção do título de Bacharel(a) em Biomedicina.

Área de atuação: Genética

Orientadoras: Profa. Dra. Vera Lúcia da Silva Valente Gaiesky e Profa. Dra. Cláudia Mieko Mizutani

Porto Alegre
Julho/2013

“We can only see a short distance ahead, but we can see plenty there that needs to be done.”

Alan Turing

"When you're curious, you find lots of interesting things to do. And one thing it takes to accomplish something is courage."

Walt Disney

Agradecimentos

Aos meus pais, João Cláudio Ambrosi e Elizabeth Ribeiro Ambrosi, que me ensinaram a sonhar, acreditar, batalhar e conquistar.

Ao meu companheiro, Vinícius Ambrosi, que me acalma e me faz sorrir sempre.

Aos meus amigos, Patrícia Renck Nunes, Mariana Schünemann e Wagner de Paula Nunes, brilhantes em suas peculiaridades, que me ensinaram que a amizade ocorre naturalmente quando agimos de acordo com o que somos e acreditamos. A convivência com vocês fez de mim uma pessoa melhor e mais feliz.

Aos professores da UFRGS, que me mostraram não apenas o conteúdo, mas as ferramentas de aprendizado. Em especial, agradeço aos professores Guido Lenz e Denise Maria Zancan, que considero meus mentores e grandes exemplos de profissionalidade. Agradeço também a todos os professores que me receberam em seus laboratórios de pesquisa, pois lá aprendi a importância da diversidade, de pessoas e de idéias, para a produção de ciência de qualidade.

Aos meus orientadores da Case Western Reserve University, Cláudia Mieko Mizutani e Hillel Chiel, por depositarem toda sua confiança em mim, e me ensinarem que o medo de errar deve estar sempre à espreita, mas não deve nos impedir de tentar.

À minha orientadora da UFRGS, Vera Lúcia da Silva Valente Gaiesky, por me acolher em seu laboratório e se dedicar a ajudar num projeto um tanto distante de sua linha de pesquisa.

Ao time do laboratório da CWRU, especialmente ao Rui Sousa Neves, ao Sebastian Chahda, e ao Joe Schinaman, que proporcionaram um ambiente de trabalho e aprendizado como nenhum outro.

Aos membros da banca, Elgion Lucio da Silva Loreto, Leonardo Gregory Brunnet e José Claudio Fonseca Moreira, por revisarem com carinho meu trabalho.

À UFRGS e ao programa Ciência sem Fronteiras da CAPES, por abrirem portas que talvez eu sozinha não conseguisse abrir.

Índice

Índice de Figuras e Tabelas

Introdução

Artigo

Material suplementar do artigo

Lista de Abreviaturas

AEL – after egg laying (após postura dos ovos)

bus – *Drosophila busckii*

Dl – Dorsal

Dpp – Decapentaplegic

GFP – Green Fluorescent Protein (proteína fluorescente verde)

Ind – Intermediate neuroblasts defective

mel – *Drosophila melanogaster*

Msh – Muscle segment homeobox

nc – nuclear cycle (ciclo nuclear)

PMad – Phosphorylated Mothers against Dpp

sec – *Drosophila sechellia*

sim – *Drosophila simulans*

TollR – Toll receptors (receptores Toll)

Tkv – Thickveins receptor

Vnd – Ventral nervous system defective

wt – wild-type (tipo selvagem)

Resumo

A padronização ao longo do eixo dorso-ventral (DV) embrionário resulta na formação das camadas germinativas e representa um dos processos de desenvolvimento mais conservados entre os animais bilaterais. Muitos dos princípios básicos desse sistema de coordenadas foram inicialmente descobertos em *Drosophila melanogaster,* cujos mecanismos de formação de gradientes morfogenéticos e regulação da expressão gênica são hoje conhecidos em grande detalhe genético e bioquímico. No entanto, uma pergunta fundamental pouco investigada é como este sistema se adapta a alterações de tamanho de embriões em espécies distintas. Estudos clássicos de neuroanatomia demonstraram há mais de vinte anos que insetos divergentes possuem o neuroectoderma com tamanho conservado, apesar de grandes diferenças de tamanho embrionário. Mais recentemente, verificou-se que o escalonamento das camadas germinativas de espécies relacionadas de *Drosophila* ocorre de forma desigual, com grande variação do mesoderma devido a mudanças na distribuição do gradiente nuclear do morfógeno Dorsal/NF-κB, que controla a padronização do eixo DV em *Drosophila*. Neste trabalho, investigamos o mecanismo de escalonamento do gradiente de Dorsal em resposta a mudanças no tamanho do embrião e na cinética da rede bioquímica envolvida na formação do gradiente. Para isto, utilizamos uma combinação de métodos experimentais e computacionais. Primeiramente, nós reproduzimos um modelo matemático da formação do gradiente de Dorsal e utilizamos resultados experimentais de mutantes e espécies-irmãs de *Drosophila* para melhorar o modelo. Nosso modelo modificado foi capaz de reproduzir os gradientes experimentais após ajustes no tamanho do embrião, no número de núcleos ao longo do eixo DV, e em parâmetros envolvidos na cinética da formação do complexo Dorsal-Cactus/Iκ-B. Nossos resultados sugerem que a distribuição do gradiente de Dorsal é mais sensível à geometria e à organização dos núcleos do embrião do que ao volume e ao número total de núcleos do embrião. Adicionalmente, nossos resultados sugerem que mudanças na amplitude do gradiente de Toll, que regula a translocação nuclear de Dorsal, poderiam ser responsáveis pelas diferentes formas do gradiente Dorsal observadas em espécies-irmãs de *Drosophila*.

Palavras-chave: *Drosophila*, Dorsal gradient, Mathematica, scaling

Introdução

Desenvolvimento embrionário de *Drosophila*

Drosophila melanogaster, também conhecida como mosca-da-fruta, é um modelo animal bem estabelecido de desenvolvimento metazoário. Com apenas 4 cromossomos, o genoma sequenciado, um ciclo de vida curto, e diversas ferramentas genéticas bem descritas, *D. melanogaster* tornou-se um modelo essencial para a descoberta de novos processos da biologia do desenvolvimento. Mais recentemente, com a descoberta e sequenciamento genômico de novas espécies relacionadas, o estudo comparativo de Drosophilideos passou a contribuir de forma significativa para o campo da biologia evolutiva do desenvolvimento, ou Evo-Devo.

Uma das perguntas-chave na biologia do desenvolvimento é como é determinado o destino de cada célula no embrião. A área de Evo-Devo estende esta pergunta e preocupa-se em entender como processos evolutivos atuam para o aparecimento de novas formas embrionárias que são vistas em espécies distintas. Neste trabalho, nós investigamos o problema de escalonamento das camadas germinativas e regulação da expressão gênica em espécies próximas e distantes de *D. melanogaster* que variam em tamanho de embrião.

Padrões de expressão gênica determinam a diferenciação de células embrionárias

Durante o desenvolvimento, células vizinhas inicialmente indiferenciadas são capazes de adquirir especializações distintas a partir da leitura de gradientes morfogenéticos. Morfógenos são *compostos geradores de forma*, isto é, eles ativam a expressão de genes-alvo de maneira dose-dependente, induzindo diferenciação de diversas células e portanto fornecendo às células informação posicional. Desta forma, determinados limiares de concentração são interpretados como um sinal para ligar ou desligar a expressão de certo grupo de genes. A expressão gênica diferencial é a responsável por resultar em desfechos distintos quanto à diferenciação celular [1].

Em *Drosophila*, o gradiente do morfógeno e fator de transcrição Dorsal/NF-κB determina as três principais camadas germinativas formadas ao longo do eixo dorso-ventral (DV): mesoderma, neuroectoderma e ectoderma, estabelecidos na região ventral, lateral e dorsal do embrião, respectivamente. O mesoderma presuntivo é estabelecido com altos níveis de Dorsal nuclear, o que

só ocorre próximo da linha ventral média do embrião. Essa camada germinativa expressa *snail* (*sna*), um gene-alvo de Dorsal cujo promotor tem baixa afinidade por Dorsal [2]. O neuroectoderma é formado nas regiões laterais, onde há níveis intermediários de Dorsal dentro do núcleo capazes de ativar genes neuroectodermais tais como *short gastrulation (sog)*. Por fim, o ectoderma é organizado com altos níveis de Decapentaplegic (Dpp/BMP) e baixos níveis de Dorsal, o que ocorre próximo da linha dorsal média. Além da expressão de *dpp*, pode-se observar a expressão de *rhomboid*, *zen* e *race* na parte dorsal do embrião, que são genes-alvo de Dpp [3].

O domínio neuroectodermal é conservado em Drosofilídios

Essa camada germinativa apresenta grande robustez no que diz respeito ao número, posição e destino das células progenitoras neurais (neuroblastos). Os números e tipos de neuroblastos são conservados em diferentes espécies de *Drosophila* e outros insetos, como louva-deus, gafanhotos e mariposas, o que reflete a importância fundamental da manutenção da arquitetura celular do neuroectoderma para a *fitness* (adaptabilidade) do organismo [4-6].

Essa propriedade especial de conservação do tamanho do neuroectoderma levanta questionamentos relacionados ao escalonamento deste tecido. *A priori*, esperaria-se que espécies com diferentes tamanhos e diferente número total de células apresentassem mudanças proporcionais em todos os seus tecidos. Contudo, como visto acima, isso não ocorre no sistema nervoso de diversos insetos. Em estudos anteriores, nosso laboratório demonstrou que as espécies-irmãs *Drosophila melanogaster*, *D. simulans* e *D. sechellia*, assim como a espécie mais distante *D. busckii*, que possuem embriões de diferentes tamanhos, apresentam uma mesma subdivisão de domínios de expressão gênica no neuroectoderma (Figura 1). Esses domínios conservados do neuroectoderma podem ser observados pela análise de três genes que subdividem o neuroectoderma dorsoventralmente, nomeadamente, *ventral nervous system defective* (*vnd*), *intermediate neuroblasts defective* (*ind*) e *muscle segment homeobox* (*msh*). Por outro lado, o mesoderma destas espécies apresenta grandes variações em tamanho, indicando que o escalonamento dos tecidos embrionários durante a evolução dá-se de forma desigual [7].

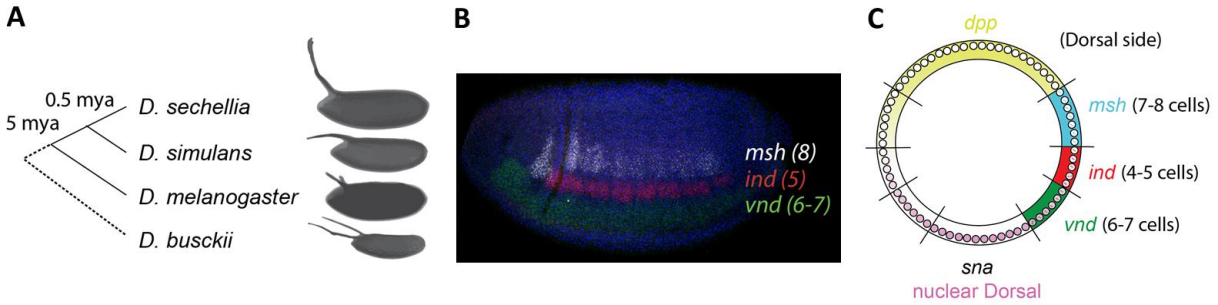


Figura 1. Diferentes espécies de *Drosophila* apresentam um neuroectoderma conservado. (A) Árvore filogenética de espécies irmãs de *Drosophila melanogaster* [8], acompanhada de imagem dos ovos das respectivas espécies. (B) Vista lateral de um embrião de *D. sechellia*, mostrando os domínios conservados de *vnd*, *ind* e *msh*. (C) Esquema de um corte transversal de embrião de *Drosophila*, ilustrando o gradiente DV de Dorsal nuclear (rosa) e de atividade de Dpp (amarelo). Os subdomínios do neuroectoderma estão representados em apenas um dos lados do embrião e o mesoderma presuntivo é indicado pela expressão de *sna*. Mya: million years ago. Os dados experimentais mostrados aqui foram produzidos por Chahda *et al*.

Buscando entender melhor os mecanismos por trás do dimensionamento das camadas germinativas dorsoventrais, neste trabalho nós estudamos os mecanismos que levam ao escalonamento do gradiente de Dorsal. A seguir, fazemos uma breve introdução sobre o desenvolvimento inicial do embrião de *Drosophila* e a formação do gradiente de Dorsal.

O embrião de Drosophila *se desenvolve como um sincício até a gastrulação*

Logo após a fertilização, ocorre a clivagem, durante a qual os núcleos do embrião de *Drosophila* se dividem rapidamente e sincronizadamente 13 vezes. Considerando que a embriogênese é uma fase crítica para organismos com desenvolvimento externo (como peixes, sapos e moscas), devido à vulnerabilidade do embrião a condições ambientais desfavoráveis e predadores, esses organismos desenvolveram estratégias para minimizar a duração da embriogênese. Uma dessas estratégias é a divisão nuclear não acompanhada de citoquinese. Isto é, os núcleos se dividem, mas continuam compartilhando o mesmo citoplasma, porque não há crescimento de membrana plasmática separando os núcleos [9]. Quando diversos núcleos compartilham o mesmo citoplasma, dá-se o nome de *sincício*.

O crescimento de membranas (celularização) ocorre apenas na interfase do ciclo nuclear 14, que inicia após a 13ª divisão nuclear. Cada ciclo nuclear inicia no começo da interfase e termina ao final da

mitose. A gastrulação ocorre após a celularização e caracteriza-se por uma série de dobramentos do embrião [9]. Além da divisão por ciclos nucleares, a embriogênese de *Drosophila* também é organizada em estágios, como ilustrado na Figura 2. Nós estamos interessados na formação do gradiente de Dorsal, que inicia no ciclo 10, até o começo da celularização, que ocorre durante o ciclo 14. Esse intervalo do da embriogênesis de *Drosophila* é caracterizado como estágio 4.

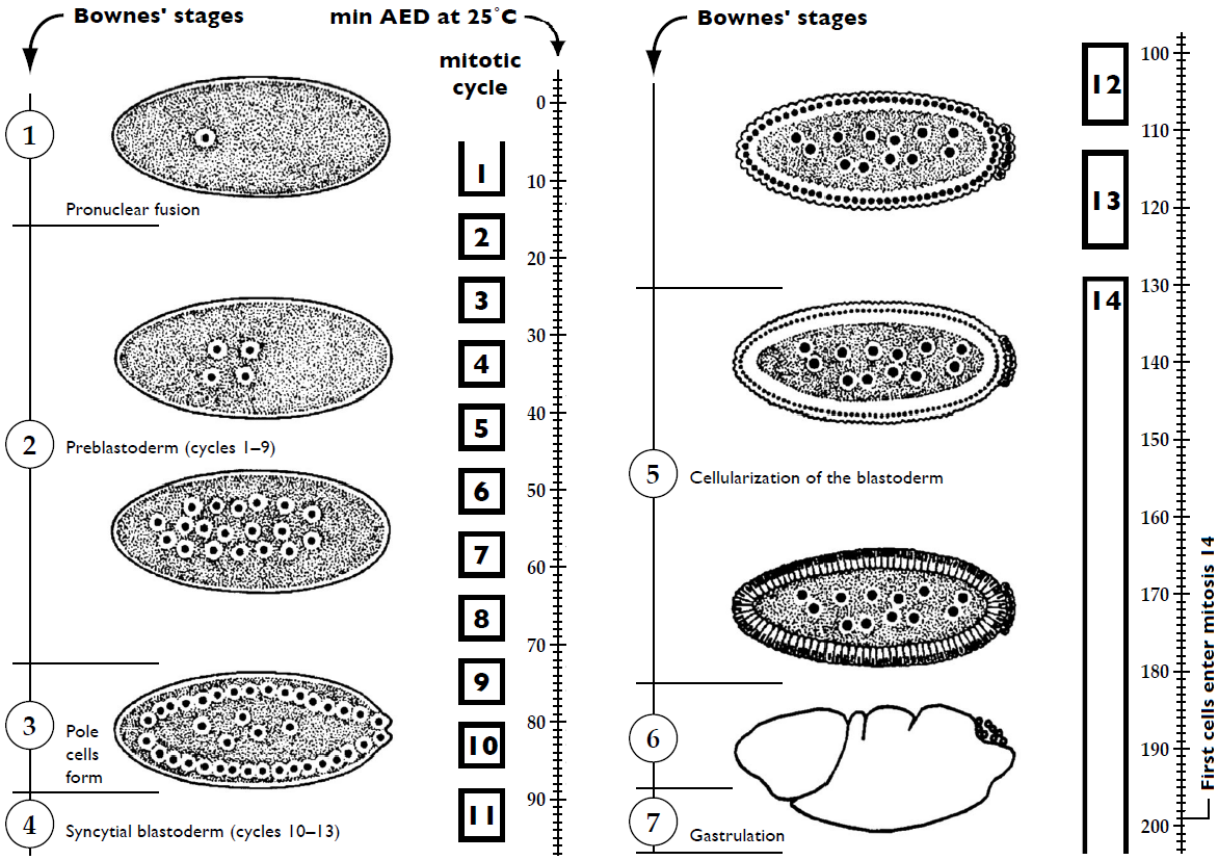


Figura 2. Linha do tempo dos primeiros 200 min de desenvolvimento de *Drosophila melanogaster*, representada em duas colunas. Desenhos representam um corte sagital do embrião (pólo anterior na esquerda e lado ventral para baixo), com exceção do último, que mostra a superfície lateral do embrião. Os retângulos numerados representam as intérfases de cada ciclo nuclear e os intervalos entre os retângulos representam a duração das mitoses. O eixo da esquerda representa os estágios do desenvolvimento como descritos por Bownes, enquanto o eixo da direita representa minutos após a deposição do ovo (AED – after egg deposition). Diferentemente dos 13 primeiros ciclos nucleares, as divisões mitóticas do ciclo 14 não são sincronizadas e começam a acontecer durante a gastrulação. Figura modificada de [10].

Fatores maternos e zigóticos determinam os os eixos embrionários, por meio de gradientes morfogenéticos

Além da clivagem sincicial, embriões de *Drosophila* empregam outra estratégia para agilizar o desenvolvimento embrionário: a deposição de proteínas e RNAs maternos. Na fase inicial da embriogênese, não há transcrição gênica, apenas tradução de mRNAs depositados maternalmente [11]. Portanto, cascatas de sinalização inteiras são herdadas do desenvolvimento do oócito. Mais de 10 fatores maternais que participam da formação do gradiente maternal de Dorsal são depositados dentro e fora do oócito. A região externa ao oócito é chamada espaço perivitelino, que é envolto pela membrana perivitelina. Essa memebrana é formada por células foliculares do ovário materno e é envolta pelo córion, que constitui a casca do ovo de *Drosophila*.

A proteína Dorsal é depositada homogeneamente no embrião, mas sua translocação nuclear é limitada pela ativação espaço-dependente de receptores Toll (TollR) inseridos na membrana plasmática do oócito. Na ausência de TollR ativados, Dorsal é confinado ao citoplasma pela interação com Cactus, um homólogo do inibidor de NF-κB, IκB. Por sua vez, o gradiente de ativação de Toll é determinado pela deposição assimétrica do mRNA de *pipe* no espaço perivitelino, e essa deposição assimétrica se dá pela falta de ativação de Epidermal Growth Factor Receptors (EGFRs) na região ventral do ovo. A transcrição de *pipe* resulta na ativação de uma cascata de proteases, incluindo os produtos dos genes *easter*, *snake* e *gastrulation defective*, que produzem a forma ativa de *späztle*, que por sua vez ativa os receptores Toll. Os receptores Toll ativados induzem a degradação de Cactus, libertando Dorsal para a translocação nuclear (Figura 3). Dessa forma, a concentração nuclear de Dorsal atinge seus níveis máximos na linha ventral média do embrião, e decresce em direção à linha dorsal média [12].

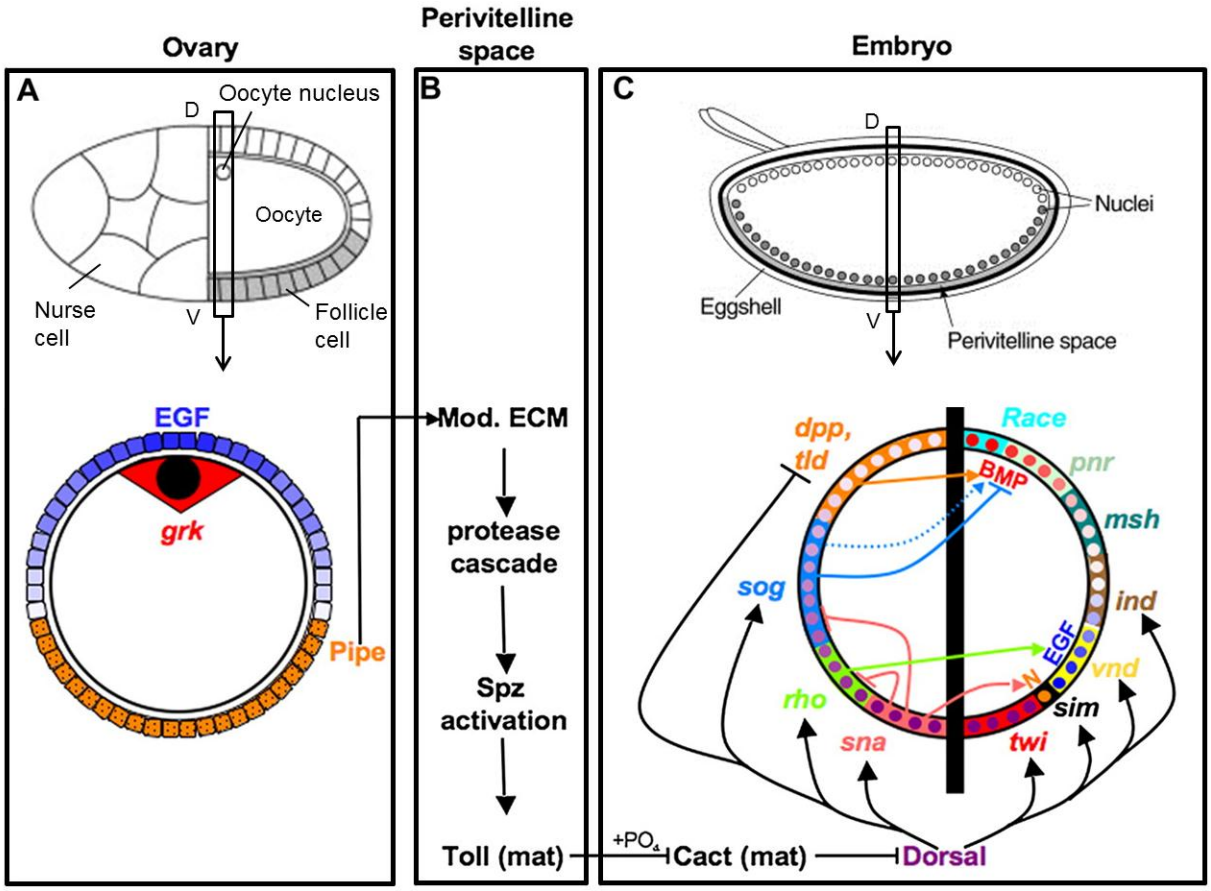


Figura 3. Padronização do gradiente DV de *Drosophila*. Em cima: vistas laterais do oócito e do embrião modificadas de LeMosy *et al* [13]. Em baixo: cortes sagitais do oócito e do embrião modificados de Lynch *et al* [14]. (A) No ovário, o oócito é envolto por células foliculares e seu núcleo se localiza na região dorsal (D). A produção do mRNA de *grk* (vermelho) ao redor do núcleo do oócito (círculo preto) resulta na ativação da via do fator de crescimento epidermal (EGF) nas células foliculares dorsais (quadrados azuis), restringindo a expressão de *pipe* às células foliculares ventrais (quadrados laranjas). (B) No espaço perivitelino, que fica entre as células foliculares e a membrana plasmática do futuro embrião, componentes da matriz extracellular (ECM, extracellular matrix) são modificados por *pipe*, levando à ativação local da cascata de proteases que ativa a proteína Spätzle (Spz). Spz ativado se liga aos receptores Toll, ativando-os. (C) No embrião, a ativação de Toll leva à degradação de Cactus, permitindo a translocação nuclear do Dorsal (roxo). Diversos genes-alvo são regulados por diferentes limiares de sinalização por Dorsal. Setas indicam ativação, e traços indicam inibição.

O gradiente de Dorsal começa a se formar apenas a partir do ciclo nuclear 10, quando os núcleos migram para a periferia do embrião. Isso ocorre porque a sinalização Toll afeta apenas a região mais próxima da camada perivitelina.

Além de Dorsal, outro morfógeno é empregado no estabelecimento do eixo DV, Dpp. Diferentemente de Dorsal, Dpp é expresso zigoticamente, ao invés de ser depositado pela mãe, e é secretado para o espaço perivitelino. Além disso, a produção desse morfógeno é limitada pela atividade de Dorsal, isto é, Dpp é expresso somente nas células expostas a baixas concentrações de Dorsal [15]. O gradiente de Dpp controla o desenvolvimento do ectoderma, que fica na parte mais dorsal do embrião, e regula as bordas do neuoectoderma junto com Dorsal.

Os genes responsáveis pelo estabelecimento do eixo dorso-ventral (DV) em *Drosophila* possuem homólogos em mamíferos que também estão envolvidos no controle da proliferação e diferenciação celular (Tabela 1) [16]. A ativação de NF-κB, por exemplo, é responsável pela resposta imune inata em eucariotos, e também possui papel importante na inflamação, proliferação e apoptose. Em *Drosophila*, além de desempenhar seu papel ancestral da resposta imune, essa via também é usada para determinar limiares de expressão gênica ao longo do eixo DV embrionário [17].

Tabela 1. Genes de *Drosophila* envolvidos no estabelecimento do eixo DV e seus homólogos em mamíferos.

Gene em *Drosophila*	**Homólogo em mamíferos**	**Referência**
dorsal (*dl*)	Rel / NF-κB	[17]
cactus	Iκ-B	[16]
decapentaplegic (*dpp*)	BMP2/4	[18]
short gastrulation (*sog*)	Chordin	[18]
ventral nervous system defective (*vnd*)	NK-2 / Nkx2.1	[19]
intermediate neuroblasts defective (*ind*)	Gsh1/2	[20]
muscle segment homeobox (*msh*)	Msx	[20]

Modelagem matemática de sistemas biológicos

O avanço tecnológico dos últimos anos permitiu a aquisição de uma quantidade cada vez maior de dados experimentais genéticos e bioquímicos do desenvolvimento, porém os métodos de análise convencionais não são suficientes para integrar todos os resultados disponíveis. Por exemplo, diversos estudos caracterizaram as proteínas envolvidas nas divisões sinciciais de *Drosophila*, mas o exato mecanismo molecular responsável pela dinâmica observada experimentalmente não pode ser testada diretamente. Em especial, muitas vezes não é possível excluir totalmente a influência de certos fatores presentes no sistema. Uma estratégia que pode ser empregada nesses casos é a modelagem matemática.

O modelo matemático pode ser usado para testar hipóteses e fazer predições que possam ser validadas experimentalmente. Em relação às divisões sinciciais de *Drosophila*, por exemplo, Calzone *et al* [21] desenvolveu um modelo matemático para testar a hipótese de que a degradação local de certa proteína depositada maternalmente (Ciclina B) seria suficiente para explicar diversas características do sistema que até então não tinham sido explicadas. Além de servir como um filtro de hipóteses viáveis, os resultados de Calzone *et al* apontam para a existência de certas relações entre as proteínas simuladas, direcionando estudos futuros.

Modelos devem ser usados para testar suficiência, mas não necessidade

A distinção entre suficiência e necessidade é fundamental na discussão de mecanismos biológicos responsáveis por um dado fenômeno. É importante ressaltar que resultados advindos de modelos matemáticos são incapazes de testar necessidade. Isso se dá porque o material de análise dos modelos foi programado, isto é, produzido por humanos, e não derivado diretamente do sistema de interesse. Mesmo se um modelo matemático sugerir fortemente que o mecanismo A é necessário para gerar tal fenômeno, é possível que a estrutura do modelo, ou seus parâmetros, ou as suposições usadas para construir o modelo estejam incorretas, e que ajustes no modelo resultem em diferentes resultados.

Por outro lado, modelos matemáticos são ideais para testar suficiência. Ao contrário das abordagens experimentais, que apresentam limitações técnicas e nem sempre permitem certas manipulações que seriam fundamentais para testar uma hipótese, absolutamente tudo no modelo pode ser alterado. A interação entre duas proteínas pode ser removida, a produção de uma proteína pode ser aumentada, e a difusão de certo componente pode ser restringida, por exemplo. Particularmente neste trabalho, nós usamos um modelo matemático para testar se mudanças em certos parâmetros, como tamanho do embrião, densidade nuclear e/ou raio nuclear seriam suficientes para reproduzir os gradientes de Dorsal observados experimentalmente.

Tradução de redes bioquímicas em sistemas de equações diferenciais

Existem basicamente três tipos de modelagens matemáticas aplicadas à biologia: modelos termodinâmicos, modelos booleanos e modelos com equações diferenciais. O primeiro é usado principalmente para estudar a integração de fatores cis-regulatórios na expressão gênica, como a interação de fatores de transcrição com sítios específicos do DNA. O segundo é usado para

desenvolver modelos qualitativos e simples, com operadores lógicos do tipo “and”, “or” e “not”. Já o último é usado para descrever a dinâmica espaço-temporal de um sistema composto por moléculas como proteínas e mRNA, e foi o tipo de modelo usado neste trabalho [22-23].

Equações diferenciais nada mais são do que equações que incluem derivadas, isto é, taxas de variação de uma variável em relação à outra variável, em geral o tempo. Essas equações são escritas a partir das reações químicas que se deseja simular. Resumidamente, cada equação representa a taxa de variação dos níveis de certa molécula em função dos outros componentes do sistema, como outras proteínas, taxas de difusão, de produção e de degradação. Para ilustrar a tradução de reações químicas para equações diferenciais, vou descrever a seguir um sistema fictício. Por exemplo, duas proteínas A e B, que não são produzidas nem degradadas, se combinam para produzir o dímero C reversivelmente. Por sua vez, o complexo C é degradado numa reação de primeira ordem. As seguintes equações químicas podem ser escritas:

$$A + B \overset{Kb}{\rightarrow} C$$

$$C \overset{Kd}{\rightarrow} A + B$$

$$C \overset{Kded}{\longrightarrow} Degradação$$

Nas equações acima, Kb representa a constante de associação entre A e B, Kd representa a constante de dissociação e Kdeg representa a degradação de C. As equações diferenciais, representado a variação temporal da quantidade das proteínas A, B e C seriam:

$$\frac{dA}{dt} = Kb\ B - Kd\ C$$

$$\frac{dB}{dt} = Kb\ A - Kd\ C$$

$$\frac{dC}{dt} = Kb\ A\ B - Kd\ C - Kdeg\ C$$

As equações diferenciais revelam grande parte dos pressupostos utilizados na modelagem matemática. No exemplo acima, a falta de termos relacionados à difusão revela que a distribuição das proteínas foi considerada homogênea no sistema. Quando maior o número de equações e termos dentro das equações, maior a exigência computacional do modelo. Devido a isso, modelos

geralmente incluem simplificações do sistema e em geral assume-se que os processos não simulados não afetam significativamente o resultado final.

O uso de equações diferenciais permite uma análise quantitativa de todas as espécies químicas simuladas. Em geral, as equações são resolvidas numericamente, dados os valores de todos os parâmetros (no exemplo acima: Kb, Kd, Kdeg) e condições iniciais (no exemplo acima: concentrações iniciais de A, B e C). Claramente, em modelos relativamente complexos, muitos parâmetros são desconhecidos. Nesse caso, os parâmetros podem ser estimados usando diversos métodos, que se baseiam no teste de milhares de grupos de parâmetros, buscando o grupo que resulta na maior semelhança entre o resultado simulado e o obtido experimentalmente, ou "*fitness*". Kanodia *et al* [23], por exemplo, no seu modelo sobre a formação do gradiente de Dorsal, desenvolveu um algoritmo genético de otimização global para determinar os valores de nove parâmetros desconhecidos do seu sistema de equações.

Modelo do gradiente de Dorsal desenvolvido por Kanodia et al

Neste trabalho, nós reproduzimos e modificamos o modelo desenvolvido por Kanodia *et al* [23] para testar nossas hipóteses. Esse modelo foi desenvolvido para reproduzir a formação do gradiente de Dorsal durante as últimas divisões sinciciais de *Drosophila melanogaster*. Uma explicação razoavelmente detalhada do modelo de Kanodia foi dada no paper apresentado a seguir, em especial nos materiais suplementares.

Referências da Introdução

1. Wolpert L (1994) Positional information and pattern formation in development. Developmental genetics 15: 485–490.

2. Ip YT, Park RE, Kosman D, Yazdanbakhsh K, Levine M (1992) dorsal-twist interactions establish snail expression in the presumptive mesoderm of the *Drosophila* embryo. Genes & Development 6: 1518–1530. doi:10.1101/gad.6.8.1518.

3. Mizutani CM, Nie Q, Wan FYM, Zhang Y-T, Vilmos P, et al. (2005) Formation of the BMP activity gradient in the *Drosophila* embryo. Developmental cell 8: 915–924. doi:10.1016/j.devcel.2005.04.009.

4. Doe CQ (1992) Molecular markers for identified neuroblasts and ganglion mother cells in the *Drosophila* central nervous system. Development 116: 855–863.

5. Whitington PM (1996) Evolution of neural development in the arthropods. Seminars in Cell & Developmental Biology 7: 605–614. doi:http://dx.doi.org/10.1006/scdb.1996.0074.

6. Thomas JB, Bastiani MJ, Bate M, Goodman CS (1984) From grasshopper to *Drosophila*: a common plan for neuronal development. Nature 310: 203–207.

7. Chahda JS, Sousa-Neves R, Mizutani CM (2013) Variation in the dorsal gradient distribution is a source for modified scaling of germ layers in *Drosophila*. Current biology : CB 23: 710–716. doi:10.1016/j.cub.2013.03.031.

8. Bächli G (n.d.) TaxoDros: The Database on Taxonomy of Drosophililidae, v. 1.03, Database 2008/01. 2008. Available: http://taxodros.unizh.ch/. Accessed 8 January 2008.

9. Foe VE, Alberts BM (1983) Studies of nuclear and cytoplasmic behaviour during the five mitotic cycles that precede gastrulation in Drosophila embryogenesis. Journal of cell science 61: 31–70.

10. V. Foe's Timetable of *Drosophila* Early Development (n.d.). Available: http://celldynamics.org/celldynamics/downloads/methods/timetable.pdf.

11. Schier AF (2007) The maternal-zygotic transition: death and birth of RNAs. Science (New York, NY) 316: 406–407. doi:10.1126/science.1140693.

12. Moussian B, Roth S (2005) Dorsoventral axis formation in the *Drosophila* embryo - shaping and transducing a morphogen gradient. Current biology 15: R887–99. doi:10.1016/j.cub.2005.10.026.

13. LeMosy EK, Hong CC, Hashimoto C (1999) Signal transduction by a protease cascade. Trends in Cell Biology 9: 102–107. doi:http://dx.doi.org/10.1016/S0962-8924(98)01494-9.

14. Lynch J a, Roth S (2011) The evolution of dorsal-ventral patterning mechanisms in insects. Genes & development 25: 107–118. doi:10.1101/gad.2010711.

15. Mizutani CM, Meyer N, Roelink H, Bier E (2006) Threshold-dependent BMP-mediated repression: a model for a conserved mechanism that patterns the neuroectoderm. PLoS biology 4: e313.

16. Belvin MP, Anderson K V (1996) A conserved signaling pathway: the *Drosophila* toll-dorsal pathway. Annual review of cell and developmental biology 12: 393–416. doi:10.1146/annurev.cellbio.12.1.393.

17. Rushlow C, Warrior R (1992) The rel family of proteins. BioEssays 14: 89–95. doi:10.1002/bies.950140204.

18. van der Zee M, Stockhammer O, von Levetzow C, Nunes da Fonseca R, Roth S (2006) Sog/Chordin is required for ventral-to-dorsal Dpp/BMP transport and head formation in a short germ insect. Proceedings of the National Academy of Sciences of the United States of America 103: 16307–16312. doi:10.1073/pnas.0605154103.

19. Stepchenko AG, Pankratova E V, Doronin S a, Gulag P V, Georgieva SG (2011) The alternative protein isoform NK2B, encoded by the vnd/NK-2 proneural gene, directly activates transcription and is expressed following the start of cells differentiation. Nucleic acids research 39: 5401–5411. doi:10.1093/nar/gkr121.

20. Cornell R a, Ohlen T V (2000) Vnd/nkx, ind/gsh, and msh/msx: conserved regulators of dorsoventral neural patterning? Current opinion in neurobiology 10: 63–71.

21. Calzone L, Thieffry D, Tyson JJ, Novak B (2007) Dynamical modeling of syncytial mitotic cycles in Drosophila embryos. Molecular systems biology 3: 131. doi:10.1038/msb4100171.

22. Ay A, Arnosti DN (2011) Mathematical modeling of gene expression: a guide for the perplexed biologist. Critical reviews in biochemistry and molecular biology 46: 137–151. doi:10.3109/10409238.2011.556597.

23. Kanodia JS, Rikhy R, Kim Y, Lund VK, DeLotto R, et al. (2009) Dynamics of the Dorsal morphogen gradient. Proceedings of the National Academy of Sciences of the United States of America 106: 21707–21712. doi:10.1073/pnas.0912395106.

Artigo Científico

O presente trabalho será apresentado nas normas da revista PLOS Computational Biology conforme as instruções que se encontram em http://www.ploscompbiol.org/static/guidelines#preparation (acessado em 06/06/2013). Para conveniência da banca, figuras, tabelas, e legendas serão inseridos no meio do texto. Quando for conveniente, dados suplementares também serão inseridos no meio do texto.

Title

Scaling and evolution of *Drosophila* embryonic germ layers – Investigation of Dorsal gradient dynamics

Authors

Priscilla Ambrosi[1,2], Juan Sebastian Chahda[2], Hillel Chiel[2], Claudia Mieko Mizutani[2]

Affiliations

[1]Department of Genetics and Molecular Biology, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

[2]Department of Biology, Case Western Reserve University, Cleveland, OH 44106, USA

Abstract

The embryonic dorsal-ventral (DV) patterning is a highly conserved developmental process in bilaterians that is responsible for establishing the primary germ layers. Many mechanisms underlying DV patterning have been described in great genetic and biochemical detail for *Drosophila melanogaster*, but fundamental questions about how this system adapts to changes in embryo size throughout evolution remain largely unexplored. Classic neuroanatomical studies have shown that divergent insects share a conserved neuroectodermal domain size, despite big discrepancies in embryo size. More recently, it was observed that the scaling of germ layers in close *Drosophila* species is uneven, and that both the size of the mesoderm and the shape of the Dorsal/NF-κB morphogen nuclear gradient, which orchestrates DV patterning, is species-specific. Here we have used a combination of experimental and *in silico* approaches to investigate the scaling process of the Dorsal gradient in response to changes in embryo size and kinetics of the biochemical network underlying the gradient formation. First, we reproduced a mathematical model of the Dorsal gradient formation and used experimental data from mutants and *Drosophila* sibling species to improve the model. With adjustments in embryo size, number of DV nuclei and kinetic parameters involved in the formation of the Dorsal-Cactus/Iκ-B complex, our modified model was able to reproduce triploid and haploid mutant gradients, as well as the gradients from *D. simulans*, *D. sechellia* and *D. busckii*. Our results suggest that the Dorsal gradient distribution is more sensitive to embryo geometry and nuclei arrangement along the axes than to overall embryo volume and total number of nuclei. Additionally, our results suggest that changes in Toll signaling amplitude, which controls Dorsal nuclear translocation, could be responsible for the distinct Dorsal gradients observed in *Drosophila* sibling species.

Introduction

One fundamental question in developmental biology that remains largely unexplored is how tissues and cell types respond to scaling. It has been known for a long time that the types and numbers of neural progenitor cells specified along the dorsal-ventral (DV) axis are highly conserved in insects that diverged for several million years [1–3]. This observation is puzzling since these species have obvious differences in total numbers of embryonic cells. Here we address the question of scaling by testing a mathematical model to predict how patterning along the DV axis responds to differences in egg size in closely related *Drosophila* species. We recently discovered that the germ layers of related Drosophilids are unequally scaled, and that the mesoderm shows significant changes in size, whereas the neuroectoderm maintains its size [4,5]. The variation in mesodermal size provides a mechanism for a new positioning of the neuroectodermal border along the DV axis of the *Drosophila* species.

In our previous work, we determined that the Dorsal/NF-κB nuclear gradient established by Toll receptor activation does not scale across the closely related species *D. melanogaster*, *D. sechellia* and *D. simulans*, and the more distantly related species *D. busckii*. In *Drosophila*, a ventral-to-dorsal concentration gradient of nuclear Dorsal (Dl) establishes the three primary germ layers, with high, intermediate and low levels of nuclear Dl being required for the specification of the mesoderm, neuroectoderm (NE) and ectoderm, respectively. Unlike other embryonic gradients, such as *bicoid*, the shape of the Dl gradient after the last syncytial division at nuclear cycle 14 is species-specific, as well as the size of the prospective mesoderm [4,6]. Here we are interested in understanding the underlying factors responsible for those differences in the Dl gradient formation in different species and ultimately understand the evolutionary mechanisms employed for germ layer specification.

The species-specific Dl gradients are likely due to a combination of biophysical and biochemical factors. Besides embryo size, the aforementioned species differ in nuclei density and radius [4]. It is likely that these species also have differences in how the Toll signaling gradient is established, since Toll/NF-κB signaling is also central to the immune response, which is a fast evolving pathway [7]. To elucidate the role of these factors on the Dl gradient shape, we combined computational and experimental approaches presented here.

We first reproduced and modified a mathematical model of the Dl gradient formation previously developed for *D. melanogaster* to test whether or not this model could reproduce the Dl gradient distributions displayed in mutant embryos with altered nuclei density and size.

These mutants have presumably wild-type (wt) embryo size and Toll gradient, but go through one extra or one less cleavage division resulting in differences in nuclei size and density. *sesame* (*ssm*) is a haploid embryo with small and tightly packed nuclei, while gynogenetic-2;gynogenetic-3 (*gyn*) is a triploid with large and loosely packed nuclei [4]. After the last syncytial division (nuclear cycle 15 for ssm and 13 for gyn), the Dl gradient shape of these mutants is significantly different from wild-type (Figure 1). We hypothesized that the mutant gradients could be reproduced by the mathematical model given that nuclei size and density were adjusted according to our experimental measurements.

Our results revealed that altered nuclei density and size are indeed sufficient to simulate the flattened ssm gradient, but they are not enough to simulate the gyn mutant gradient. Our model predicts that additional changes in embryo geometry are partly responsible for the steep gradient observed in gyn mutants. We next employed the mathematical model to predict which factors influence Dl gradient formation observed in different *Drosophila* species. We conclude that the Dl gradient shape is most sensitive to parameters that determine the distribution of Dl-Cactus along the DV axis, and that the amplitude of Toll signaling could be responsible for the species-specific Dl gradients.

Results

We hypothesized that changes in nuclei density and size are sufficient to generate the distinct Dl gradient shapes of ssm and gyn mutants. We used a mathematical model designed to reproduce the wt gradient formation as our experimental framework, because a qualitative analysis is not sufficient to predict the combined effect of changes in nuclei density and size on the Dl gradient shape. As shown in Figure 1, if we consider that all three genotypes have a similar Toll signal distribution, nuclei density affects the difference in Toll signal reading from one nucleus to the other, affecting the rate of decay of the Toll signal. Hence, a small and a large nuclei density would lead to a steep and a flat gradient, respectively. However, if we consider that all genotypes have the same total amount of Dl protein, nuclei density also affects the amount of Dl per cell compartment. Additionally, differences in nuclei size affect the rate of Dl translocation, as well as the Dl nuclear concentration. A large nucleus has more surface area available for Dl transport, but it also has a large volume, preventing nuclear Dl from reaching higher concentrations. Such nuclear features could counterbalance the effects of nuclei density. Hence, a numerical approach is necessary to determine whether or not the combined effects of nuclei density and size are sufficient to produce the distinct Dl gradient shapes observed in the mutants. Such numerical approach was based on a mathematical model developed by Kanodia *et al* [8] (hereafter referred as “Kanodia model”).

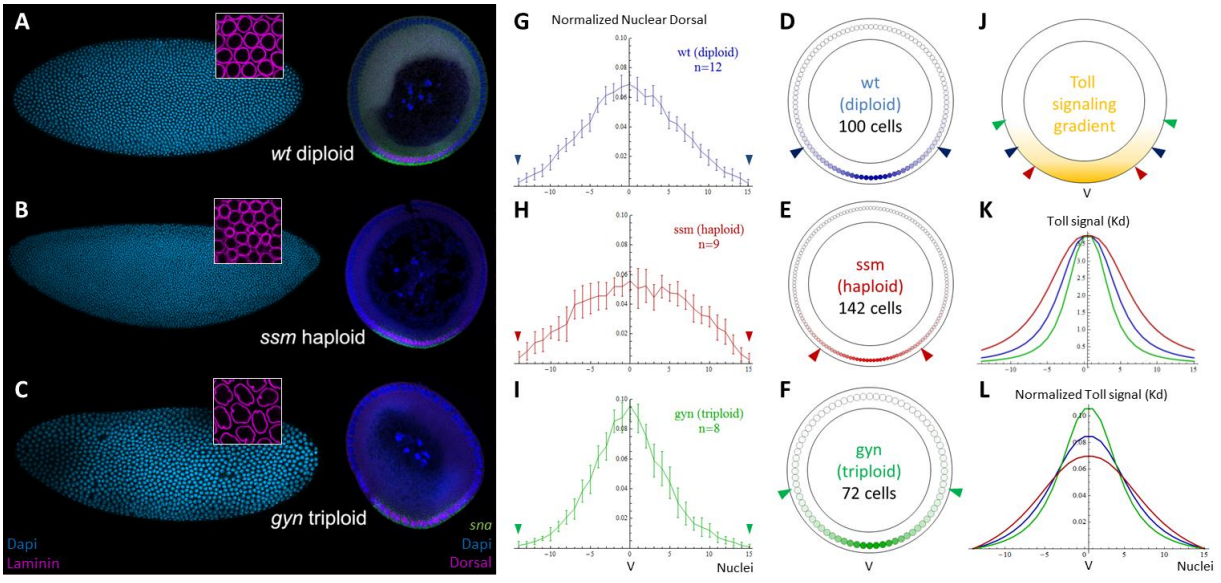
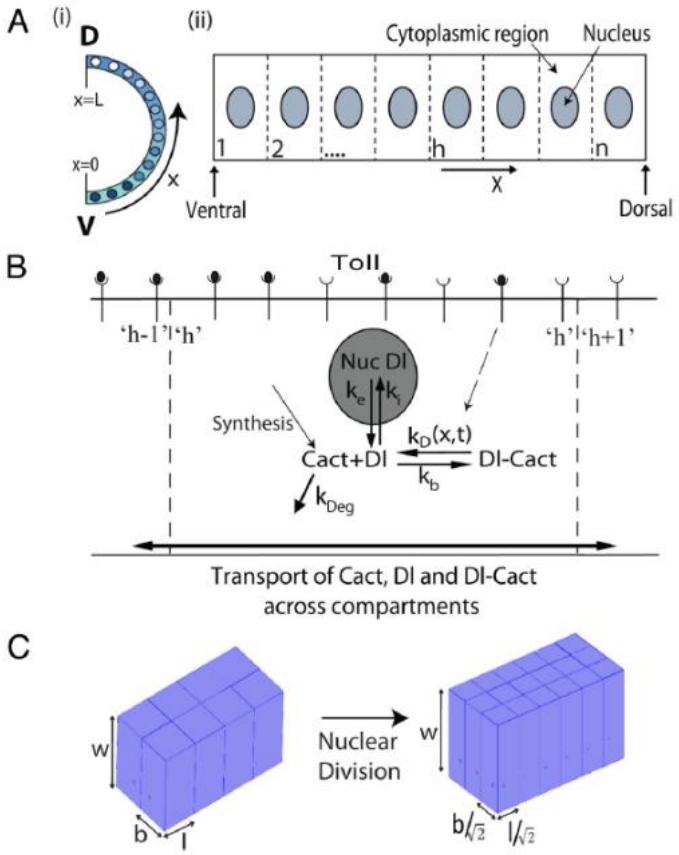


Figure 1. *D. melanogaster* mutants ssm and gyn have distinct Dl gradient shapes in comparison with wt, and nuclei density affects the rate of decay of the Toll signal at the 30 most ventral cells. (A-C) Whole-mount blastoderm embryos (left) and corresponding cross-sections (right) stained with DAPI

(blue), anti-Dl (magenta) and sna mRNA (green). Insets show increasing nuclear size and decreasing nuclei density from haploids (B) to diploids (A) to triploids (C) stained with anti-laminin (magenta). (G-I) Normalized intensity levels of nuclear Dl per nuclei (mean±SD), considering the 30 most ventral nuclei only (the ventral midline is represented by V at x=0). (D-J) Cross-section schemes for wt (D, blue), ssm (E, red), gyn (F, green), and a general *D. melanogaster* embryo (J) representing the Toll signaling gradient. Color-coded arrowheads delimit the 30 most ventral cells for each genotype. (K) Simulated Toll signaling gradient illustrated in (D-J) and based on the equation for Kd, the space-dependent Cactus degradation constant (see Methods). As illustrated in (J), nuclei density affects the angle subtended by 30 cells in a cross-section, resulting in a larger rate of Toll signal decay for gyn and a smaller rate for ssm. (L) Normalized Toll signaling gradients, emphasizing the relationship between the simulated Toll signaling gradient and experimental Dl gradients. Figures (A-I) were modified from Chahda *et al* [4].

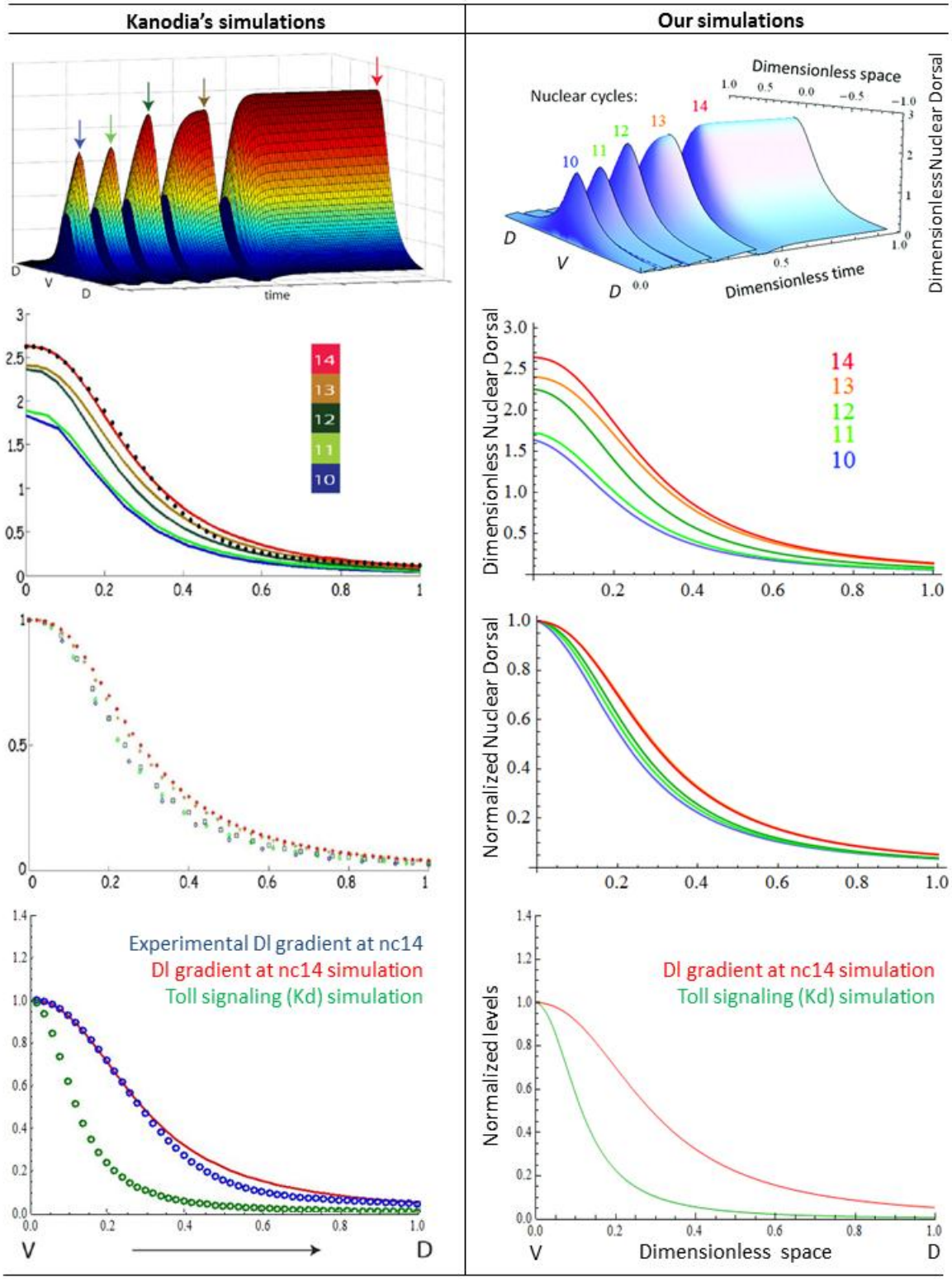
Reconstruction of the Kanodia model reproduces original simulations of Dl gradient dynamics

Key features of the Dl gradient formation from nuclear cycle (nc) 10 to 14 in a DV cross-section of the embryo were modeled by Kanodia *et al* [8]. Essentially, differential equations were derived from mass balance equations involving Dl and Cactus/IκB (Supplementary Figure 1, Supplementary Methods) and then numerically solved for the time intervals determined by Foe and Alberts [9] using globally optimized parameters. A DV cross-section of the syncytial blastoderm was simplified as a string of cuboid cellular compartments containing a spherical nucleus and its surrounding cytoplasm, while the Toll signaling gradient was represented by a space-dependent reaction rate constant for Dl-Cactus dissociation (Kd).



Supplementary Figure 1. Kanodia model rationale. (A) The embryo is modeled as a single string of n cuboid cellular compartments (represented by h), arranged in a row of length L (half-circumference of the DV cross-section). (B) Chemical reactions and transport processes considered by the model. (C) Changes in the cell compartment dimensions after a nuclear division. Note that the height of the compartments (w) is constant and that the volume halves. This figure was reproduced from Kanodia *et al* [8].

In order to employ the Kanodia model to reproduce our mutant ssm and gyn gradients, it was imperative that our reconstructed model using Mathematica could first reproduce Kanodia simulations of the wt gradient. Supplementary Figure 2 shows that our model reproduces key features of Kanodia simulations. Slight differences are probably due to software-specific algorithms used to numerically solve differential equations.



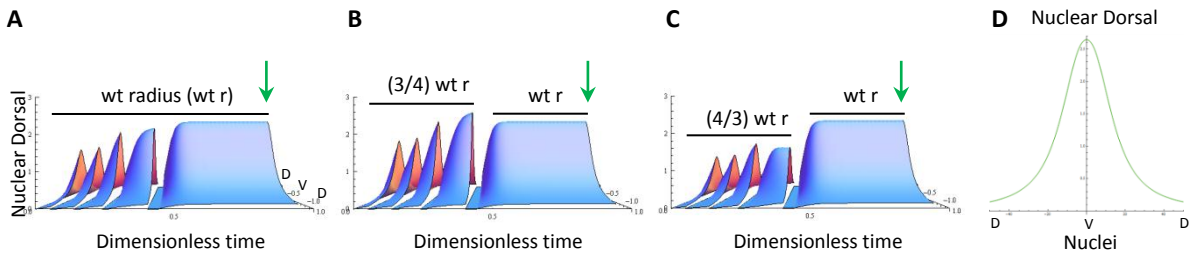
Supplementary Figure 2. Comparison between original simulations from Kanodia *et al* in Matlab [8] and our simulations from the model reproduced in Mathematica. First row: 3D plots of wt Dl gradient from nuclear cycle 10 to 14 (nc10 – nc14). Second row: amplitude of the Dl gradient at the end of each

interphase (see color code). Third row: Dl gradient at the end of each interphase (same color code) normalized as a percentage of the highest nuclear Dl level, highlighting that the shape of the Dl gradient is conserved throughout development. Last row: shape of the space-dependent parameter Kd, which represents Toll signaling (green), and shape of the Dl gradient at the end of nc14 (red). V: ventral midline; D: dorsal midline.

The shape of the gradient at the last nuclear cycle is not affected by early cycles

One of the Kanodia model findings was that the shape of the wt Dl gradient is constant throughout development, which matches quantifications of the Dl protein obtained experimentally [8]. Hence, we wanted to know if the shape of the ssm and gyn mutant gradients was altered from the onset of the Dl gradient formation, at nuclear cycle 10 (nc10), or if the shape was altered only at the last cycle.

Before doing so, we tested how changes in nuclear radius at nc10-13 affected the wt gradient at nc14. Surprisingly, the shape of the Dl gradient at early cycles does not affect the shape of the gradient at the last nuclear cycle (Supplementary Figure 3). It is unlikely that this result implies a biologically meaningful prediction, but rather that the parameter set used creates an artificially robust system. This result has two main consequences. First, it minimizes the effect of incorrect assumptions about early cycles on the shape of the gradient at the final cycle we are most interested in. Second, it discards the possibility of investigating the gradient at early cycles based on experimental data from the last cycle. In other words, knowing the shape of the gradient at the last cycle is not sufficient to infer about previous states of the gradient. This precludes our analysis of the dynamics of the mutant gradients in early cycles and compels us to focus on the last cycle only.



Supplementary Figure 3. Simulations of changes in nuclear radius and their effect on the Dl gradient shape. (A-C) Changes in nuclear radius (r) at nc10-13 affect the shape of the gradient at the respective

cycles, but not at the last nuclear cycle. (D) Superposition of the gradients at the end of the last nuclear cycle from A-C (green arrows). V: ventral midline; D: dorsal midline.

Our reconstructed model can reproduce the flattened Dl gradient obtained experimentally in ssm mutants

In our first attempt to reproduce our experimentally obtained Dl gradients from mutants, we used a representative set of parameters used by Kanodia *et al* [8] to reproduce the wt gradient, with changes in nuclei size and density only (see Methods), according to our experimental measurements (Figure 2). Some few additional parameters were also changed, especially related to early cycles. However, considering the model robustness, these additional parameter changes do not significantly affect our results (see discussion in Supplementary Methods, as well as Supplementary Table and Figure 1).

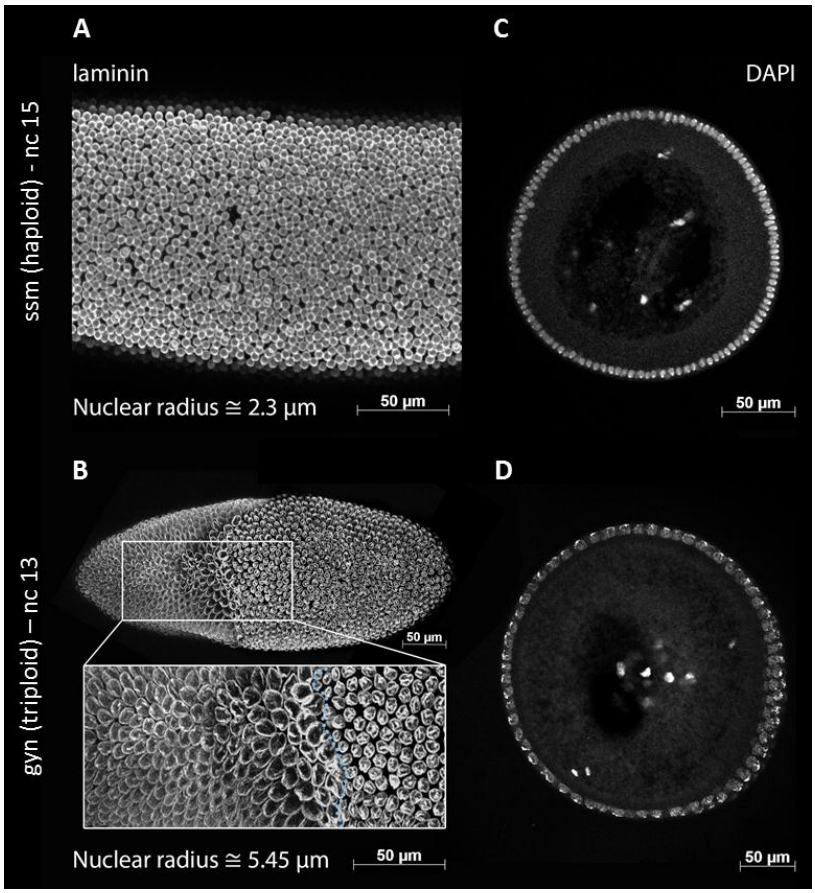


Figure 2. Nuclei radius measurements in mutants at their last nuclear cycle before cellularization. (A-B) Laminin antibody stainings for ssm (A) and gyn (B) embryos whole-mounted using glass beads to avoid compression of tissue. The triploid gyn embryo is a mosaic, whose boundary is indicated by a blue dotted line. (C-D) Hand cut ssm (C) and gyn (D) embryo cross-sections stained with DAPI.

First, the model output was normalized to match our experimental data (see Methods), which is restricted to the 30 most ventral cells instead of the entire embryonic cross section. In addition, we represent the overall shape of the Dl gradient rather that the absolute values of Dl concentration. Figure 3 compares the whole gradient to its subset containing the 30 most ventral cells, as well as illustrates the difference between normalized and non-normalized gradients. Unless otherwise noted, the normalized gradient restricted to the 30 most ventral cells is referred to as “Dl gradient”.

The simulation shown in Figure 3B suggests that ssm embryos have the highest peak of nuclear Dl concentration, while gyn embryos have the smallest peak. Hence, even though both the amount of Dl per cell compartment and the nuclear surface area available for Dl translocation are smaller in ssm, the model predicts that the smaller nuclear volume is the major determinant of the absolute concentration of nuclear Dl. In terms of Dl gradient shape, Figures 3D-F show that the model is able to reproduce the flattened haploid gradient, but not the steep triploid one, and Figure 3C shows that the simulated gyn Dl gradient has the same shape as wt.

The ability of the model to reproduce the ssm gradient but not the gyn gradient points to two non-exclusive deductions: (1) changes in nuclei density and size are sufficient to explain ssm distorted gradient, but not gyn, i.e. our hypothesis is only partially correct; and (2) the parameter set used creates a very strong artificial robustness, buffering the effect of our manipulations. To investigate if our manipulations were being buffered, we tested the individual effects of nuclei density and size on the Dl gradient shape. Figure 4 shows that either larger nuclei density or larger nuclear size result in a flattened gradient. Hence, the flattened gradient seen in ssm is mostly determined by its larger nuclei density, which overcompensates the effect of its smaller nuclei. Conversely, the effect of larger nuclei in gyn was only slightly compensated by its reduced nuclei density, resulting in a Dl gradient shape similar to wt in our simulations, rather than a steep gradient as obtained experimentally. Therefore, it is likely that some of the assumptions that apply to wt and ssm do not apply to gyn. One possibility is that one or more general parameters, such as Dl diffusion rates and Dl nuclear export rates, are different from the values employed in the model, but their effects are more significant in gyn.

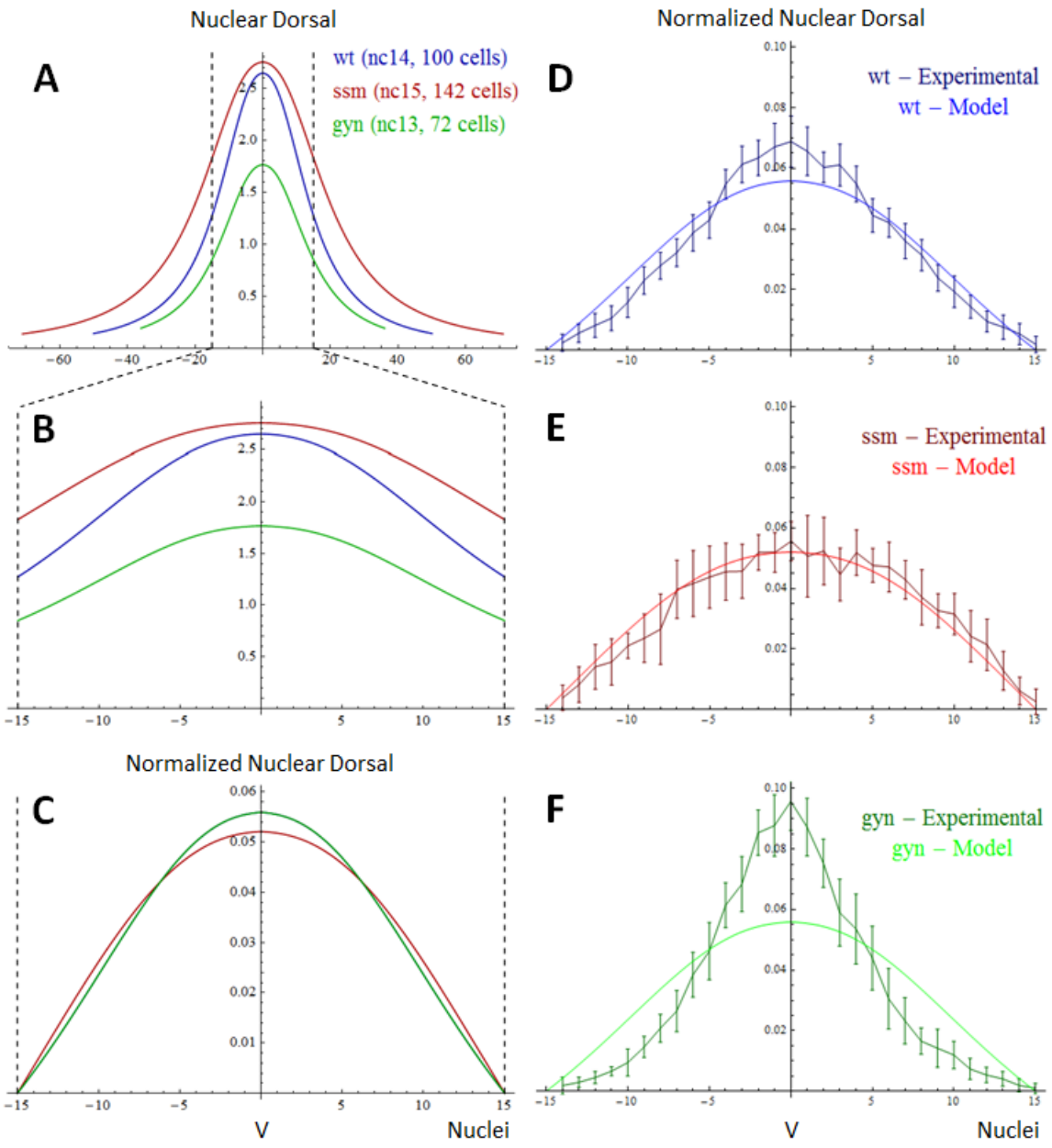


Figure 3. Comparison between experimental data and the model output. (A) Simulated nuclear Dl levels per nuclei at the last nuclear cycle of each genotype for the entire cross-section. Note that the cross-section has the same size in all genotypes, but the number of nuclei changes, due to extra or fewer nuclear cycles. (B-C) Simulated non-normalized (B) and normalized (C) Dl gradient considering the 30 most ventral cells only. In (C), the wt gradient is superimposed by gyn. (D-F) Direct comparison between experimental data and the model normalized output. The experimental data shown here (average±SD) was reproduced from Chahda *et al* [4].

We also observe that the model reproduction of our wt gradient is not completely satisfactory, which further suggests that the parameter set used could be improved. Hence, we modified the model in order to determine which biologically meaningful parameter combinations could better reproduce our experimental Dl gradients.

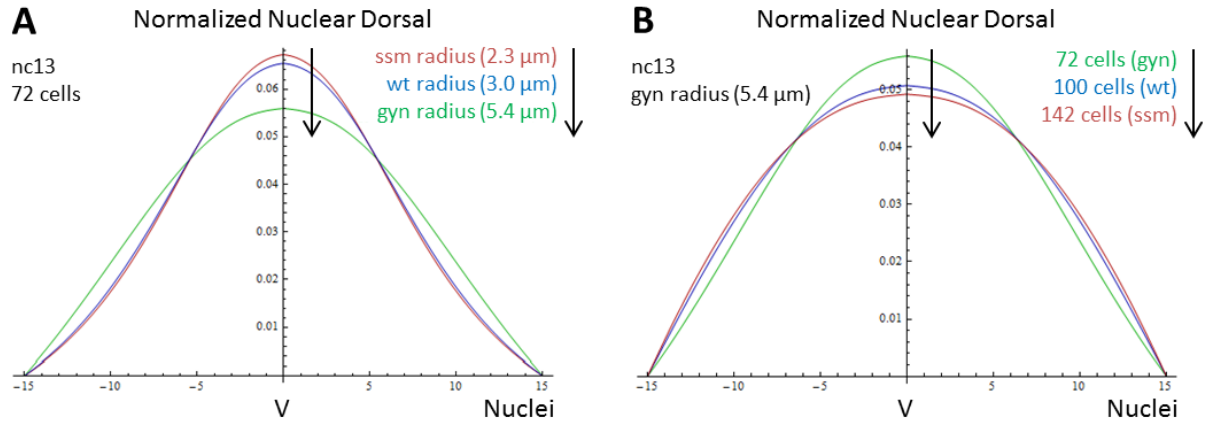


Figure 4. Individual effects of changing nuclei radius and density on the Dl gradient at nuclear cycle 13 (nc13). (A) Increasing nuclei radius and (B) density flattens the gradient, as indicated by arrows. V: ventral midline.

Refinement of the parameter values reveals that Dl diffusion and export rates, as well as embryo geometry, play major roles in the model reproduction of the Dl gradient in gyn mutants

To investigate which parameter changes allow the reproduction of the gyn gradient, we modified the original Kanodia model to increase its flexibility. Instead of the non-dimensionalized equations, we used dimensionalized equations and focused on the simulation at the last nuclear cycle (see Methods). In the original model, 9 dimensionless parameters were used, in addition to nuclei radius and density, developmental timing and cell compartment volume at nc14 (Supplementary Table 2). In our modified model, a total of 19 parameters could be manipulated (Table 1), and their effects on the Dl gradient shape could be directly analyzed (Supplementary Figure 4).

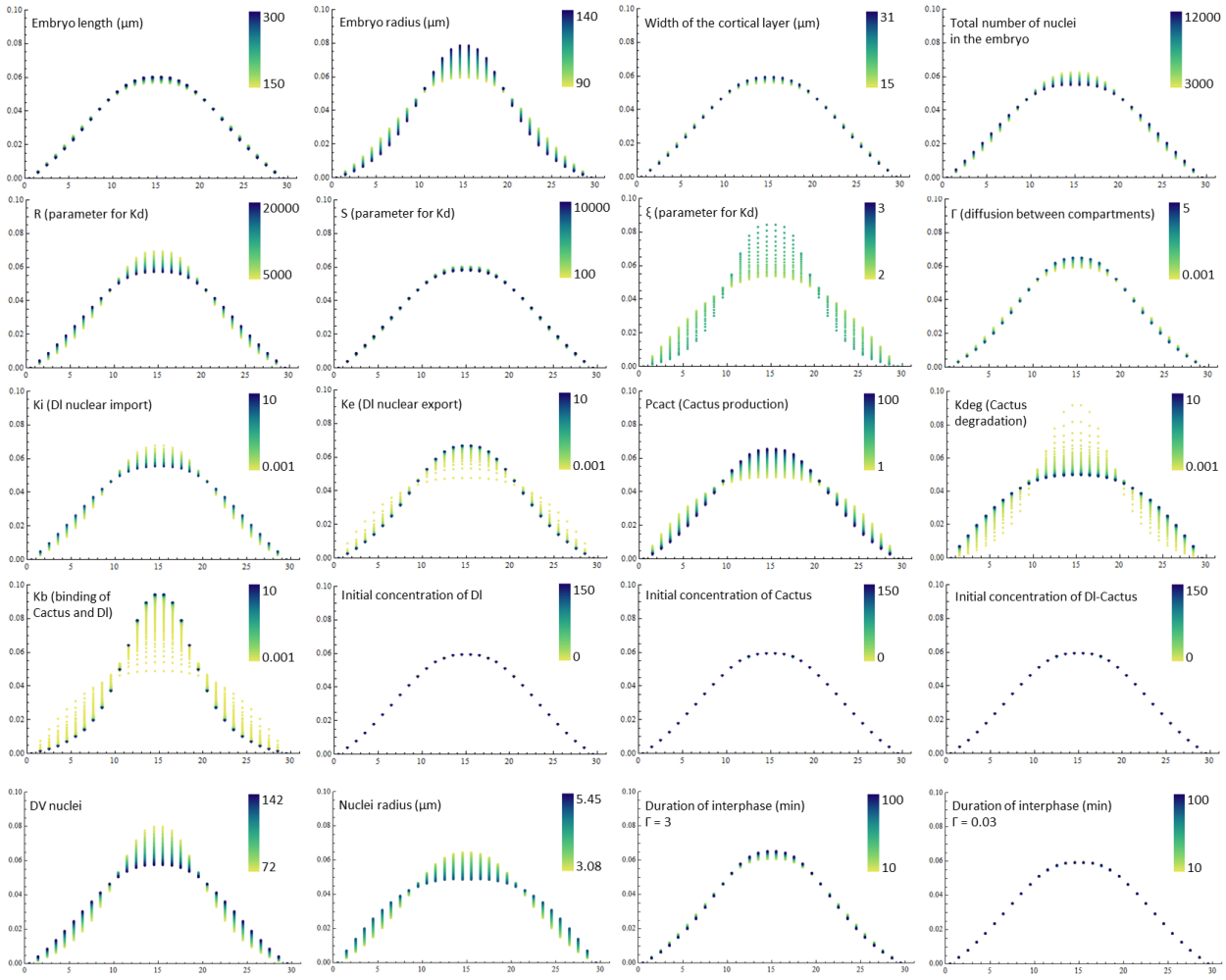
The original values of most of these 19 parameters could be estimated from the nondimensionalized parameter set, while others were determined by direct measurements and assumptions (Supplementary Table 1). The values of some parameters were unknown, due to the previous nondimensionalization. Revisions of the parameter values were performed by manually testing a combination of parameters able to reproduce both wt and

mutant gradients. To further validate the revised parameter set, we also tested the model ability to reproduce the Dl gradient of *D. melanogaster* embryos derived from dl^-/dl^+ mothers. These embryos presumably have the same wt Toll signaling and embryo size, but only half of the amount of Dl protein.

Table 1. Original and revised parameter values used for wt and mutants.

Parameter	Original values (wt)	Revised values (wt)	Embryos from dl^-/dl^+ mothers	ssm	gyn
Embryo length (μm)	245	245	245	245	245
Embryo radius (μm)	90	90	90	90	**120**
Width of cortical layer (μm)	31	31	31	31	31
Total number of nuclei in embryo	6000	6000	6000	**12000**	**3000**
R (parameter used for Kd)	15340	15000	15000	15000	15000
S (parameter used for Kd)	4052	4000	4000	4000	4000
ξ (parameter used for Kd)	2.38	**2.50**	2.50	2.50	2.50
Γ (diffusion between compartments)	0.03	**3**	3	3	3
Ki (Dl nuclear import)	1.97	**4.00**	4.00	4.00	4.00
Ke (Dl nuclear export)	0.44	0.44	0.44	0.44	0.44
Pcact (Production of Cactus)	?*	50	50	50	50
Kdeg (Cactus degradation)	1.36*	**1**	**3**	1	1
Kb (Binding of Cactus to Dl)	?*	0.03	0.03	0.03	0.03
Initial Dl/compartment (mol/L)	?**	36	**18**	36	36
Initial Dl-Cact/compartment (mol/L)	?*	30	**15**	30	30
Initial Cact/compartment (mol/L)	?*	36	36	36	36
Nuclear radius (μm)	3.08	3.08	3.08	**2.30**	**5.45**
Number of DV cell compartments	100	100	100	**142**	**72**
Duration of last nc interphase (min)	65	65	65	**55**	**86**

* In the original parameter set, the following relationship existed between these values:
Cactus = 0.93 / Kb = 1.22 * Dl-Cactus = Pcact / Kdeg
** The initial Dl concentration is thought to be similar to the initial Cactus concentration [8].
Revised parameter values are highlighted in red, while mutant-specific parameters are highlighted in orange.



Supplementary Figure 4. Individual influence of 19 parameters on the final shape of the Dl gradient. All graphs represent normalized nuclear Dl levels for the 30 most ventral cells of a cross-section at the end of nuclear cycle 14. With the exception of the parameter being manipulated, the values of the parameters used for all simulations are shown in Table 1, under revised values (wt). Note that the last two graphs show the effect of duration of interphase in two scenarios: one with low diffusion rates (Γ=0.03) and one with higher diffusion rates (Γ=3).

The analysis of embryos derived from dl^{-}/dl^{+} mothers provided valuable insights about the model parameters. These mutant embryos have a flattened Dl gradient, suggesting that at the ventral midline area, all cytoplasmic Dl is translocated into the nucleus. Hence, it is reasonable to assume that the Dl nuclear import rate (Ki) is not the limiting factor for the formation of the gradient peak. In other words, given enough Toll receptor activation and cytoplasmic Dl, peak levels of nuclear Dl can be achieved. This result motivated us to increase the value of Dl nuclear import rates, or Ki (Table 1).

Surprisingly, we found that the shape of the Dl gradient is not sensitive to the initial concentrations of Dl, Cactus and Dl-Cactus (Supplementary Figure 4). Unless the initial concentration of both Dl and Dl-Cactus is zero, in which case the Dl gradient is not formed, the shape of the Dl gradient is not affected by changes in these parameters. This suggests that the gradient shape observed in dl^-/dl^+ embryos is caused by additional parameter changes besides initial Dl concentration. Several studies report that Cactus is stabilized in the presence of Dl and that Cactus levels are reduced if Dl levels are diminished [10–12]. Based on this information, we tested if changes in the rate of Cactus degradation (Kdeg) were able to reproduce the mutant gradient. We found that doubling the wt Kdeg value was not enough to completely reproduce the dl^-/dl^+ mutant flattened gradient (data not shown). This finding suggests that the relationship between Dl amount and Cactus stabilization is not linear and probably involves cooperativity. Indeed, Dl is reported to form dimers, and the Dl-Cactus complex is generally formed by one unit of Cactus bound to two units of Dl [10,12]. Thus, we tested a higher increase of Kdeg by 3 times and under this condition the model could correctly reproduce the Dl gradient from dl^-/dl^+ embryos (Figure 5).

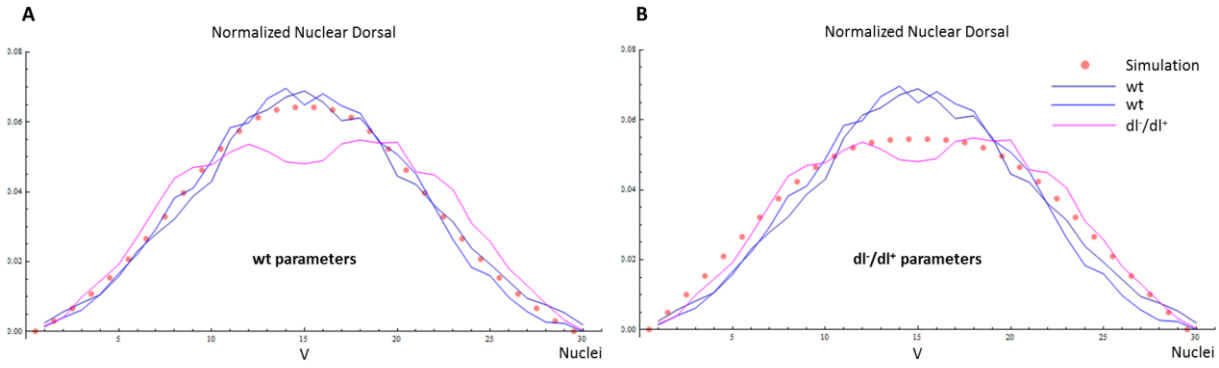


Figure 5. Changes in Kdeg allow the reproduction of the Dl gradient of embryos derived from dl^-/dl^+ mothers. (A) Comparison between the wt gradient simulation (dotted pink line) using the revised wt parameters and experimental wt Dl gradients (dark blue line: n=12, blue line: n=2). (B) Comparison between the dl^-/dl^+ gradient simulation (dotted pink line) and experimental data (pink, n=2). The parameters used in the simulations are shown in Table 1. V: ventral midline. Credit of experimental Dl gradients (average): Chahda *et al* (unpublished data).

After establishing a better wt parameter set, we tested whether such parameter combinations could then generate the gyn mutant gradient, while preserving the model ability to reproduce wt, ssm and dl^-/dl^+ gradients. Despite implementing the particular parameter set for gyn using the measured nuclei radius and density obtained experimentally, our modified model still

failed to simulate the gyn gradient, which suggests that modifying additional parameters could improve the simulations for this mutant. We were able to reproduce the gyn gradient under two additional conditions: using a higher Dl diffusion rate among compartments than the diffusion rate used in the original parameter set, and using an increased embryo radius (Figure 6 and Supplementary Figure 5). The use of this higher Dl diffusion rate kept a good fit between the simulated and experimentally obtained gradients of wt, ssm, and dl^{-}/dl^{+} embryos, and in the case of wt embryos, the fit was actually improved. An almost perfect fit for the gyn gradient was obtained by increasing the embryo radius from 90 to 120 µm. However, the use of a higher embryo radius for gyn simulations is not ideal since it is not well supported by our experimental data. In fact, the main motivation to use gyn mutants to test the influence of nuclei size and density was the fact that this mutant should *a priori* have a similar egg size as the wt, and a normal DV signaling pathway. We then looked for parameter changes that could substitute an increase in embryo radius to generate a good fit for the gyn gradient.

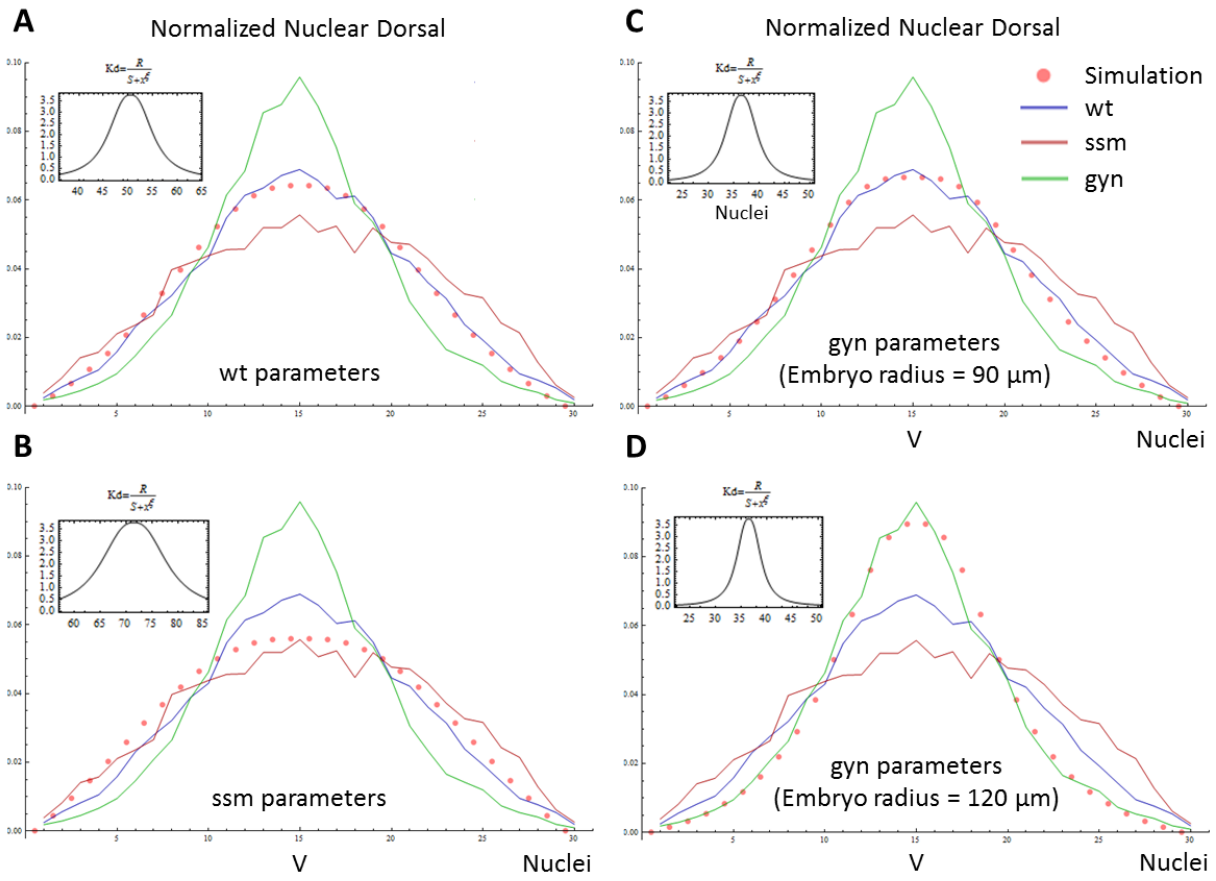
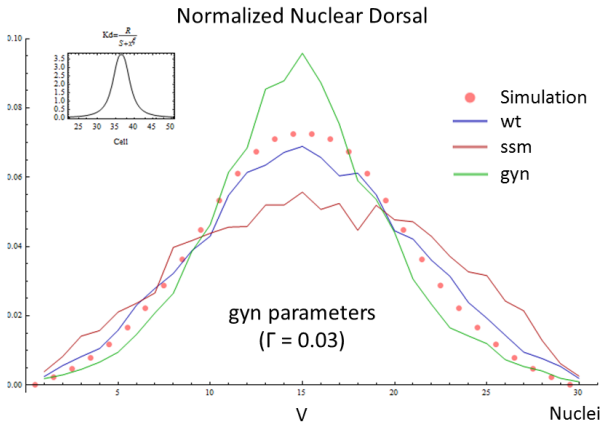


Figure 6. The modified model with the revised parameter set is able to reproduce wt and ssm gradients, and predicts that embryo geometry plays an important role on the shape of the gyn mutant gradient. (A-D) Comparison between experimental data (color-coded lines) and the simulations (pink

dotted lines) using the revised wt (A), ssm (B) and gyn (C,D) parameter sets shown in Table 1. Note that in (C) the wt value of embryo radius is used. Insets show the Toll signaling gradient (represented by Kd) for the 30 most ventral cells. V: ventral midline. Experimental Dl gradients (average) were reproduced from Chahda *et al* [4].



Supplementary Figure 5. If Dl diffusion rates are low, changes in embryo geometry in addition to nuclei density and radius are not sufficient to generate the gyn mutant gradient. The parameter set used in this simulation is shown in Table 1, but the value of Γ is the one shown in the figure (0.03 instead of 3).

By increasing the Dl nuclear export rate to 2 (an increase of roughly 4.5 times from 0.44), the fit for gyn improved significantly, requiring a smaller increase in embryo radius (from 90 to 110 µm, Figure 7). This change also did not compromise the wt and ssm simulations, and only resulted in a small increase of the Dl peak in the dl^-/dl^+ gradient simulation (data not shown). Together, an increased Dl diffusion and export rates greatly improved the gyn gradient simulation, without affecting too much the other *D. melanogaster* simulations. We believe these are reasonable assumptions to be made, since it is possible that the Dl diffusion and export rates were underestimated by using a larger Dl-GFP fusion protein than the endogenous Dl, which also contains a deletion in a nuclear export sequence. However, it remains to be tested if our much larger value for Dl diffusion of over 100 times the original value is within a reasonable range.

Despite our inability to faithfully simulate the gyn gradient using parameters that do not change embryo radius, our results nonetheless indicate that embryo geometry significantly

affects the Dl gradient shape, and is likely to play an important role in modifications seen in the other *Drosophila* species.

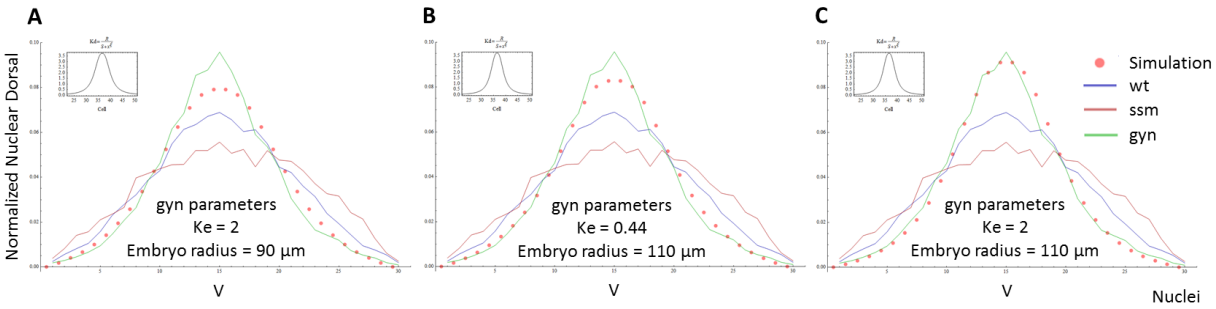


Figure 7. Changes in Dl nuclear export rates (Ke) improve gyn gradient simulations. (A-C) Simulation of the gyn gradient using higher Ke and wt embryo radius (A), low Ke and larger embryo radius (B), and a combination of higher Ke and larger embryo radius (C). Note that the simulation in (A) provides a better fit than the simulation in Figure 6C and that the simulation in (C) provides a better fit than the simulation in Figure 6D with a lower value of embryo radius.

Parameters involved in the regulation of Cactus concentration play an important role on the determination of *Drosophila* species-specific Dl gradients

Since embryo geometry, nuclei density and nuclei radius affect the shape of the Dl gradient, we wanted to test how these and other parameters act together to generate the species-specific Dl gradients observed in *D. melanogaster*, *D. busckii*, *D. simulans*, and *D. sechellia*. Table 2 shows that embryos from these species have different sizes and geometries, as well as distinct nuclei density and size. However, differently from the mutants analyzed so far, additional parameters are expected to vary from one species to the other, especially the ones related to Toll signaling, given that these species have distinct ranges of peak Toll activation levels [4]. Hence, we first tested how the physical parameters of embryo and nuclear sizes affected the Dl gradient distribution, and then we analyzed which parameter changes were sufficient to reproduce the gradients observed experimentally.

Table 2. Parameter values used to simulate the Dl gradient from *Drosophila* sibling species.

Parameter	***D. melanogaster***	***D. busckii***	***D. simulans***	***D. sechellia***
Embryo length (µm)	**241.5***	**189***	**236***	**286.5***
Embryo radius (µm)	**102.4***	**93.4***	**102.5***	**134.85***
Width of cortical layer (µm)	**23.7***	**23.4***	**30.1***	**29.6***
Total number of nuclei in embryo	6000	6000	6000	7000**
R (parameter used for Kd)	**15000**	**5000**	**75000**	**30000**
S (parameter used for Kd)	4000	4000	4000	4000
ξ (parameter used for Kd)	**2.50**	**2.70**	**2.00**	**2.35**
Γ (diffusion between compartments)	3	3	3	3
Ki (Dl nuclear import)	4.00	4.00	4.00	4.00
Ke (Dl nuclear export)	0.44	0.44	0.44	0.44
Pcact (Production of Cactus)	**50**	**150**	**10**	**25**
Kdeg (Cactus degradation)	**1**	**0.33**	**5**	**2**
Kb (Binding of Cactus to Dl)	**0.03**	**0.08**	**0.003**	**0.01**
Initial Dl/compartment (mol/L)	36	36	36	36
Initial Dl-Cact/compartment (mol/L)	30	30	30	30
Initial Cact/compartment (mol/L)	36	36	36	36
Nuclear radius (µm)	**3.08**	**2.00***	**2.75***	**3.50***
Number of DV cell compartments	**100**	**85***	**98***	**102***
Duration of last nc interphase (min)	65	65	65	65

*Experimental measurements taken from Chahda *et al* [4], which are consistent with measurements from Markow *et al* [13].
** Value based on measurements by Miles *et al* [14].
Values in red represent single parameter changes that can be made to reproduce the different species gradients. Only one of the red parameters should be changed at a given simulation and the other red parameters will be determined by the values used for *D. melanogaster*.

Figure 8A-C clearly shows that changes in embryo geometry and volume, in addition to changes in nuclei density and size, are not sufficient to reproduce the species-specific Dl gradients. We next tested if the model could predict which additional parameters might vary across species. Figure 8D-F illustrates how changes in R, one of the parameters involved in

the Toll signaling gradient, enables the reproduction of the species gradients. Interestingly, three parameters (R, Kdeg and Pcact) are analogously linked to the shape of the Dl gradient. In order to reproduce *D. busckii* gradient, the wt values of R or Kdeg can be divided by 3, or the value of Pcact can be multiplied by 3. Similarly, to reproduce *D. simulans* gradient, R or Kdeg can be multiplied by 5, or Pcact can be divided by 5. The relationship between Kdeg and Pcact was expected, because these parameters have opposing effects on the Cactus steady state concentration. R, on the other hand, controls the amplitude of Kd, i.e. the peak of Toll signaling gradient. These parameters ultimately regulate the amount and distribution of Dl-Cactus in the system (Supplementary Figure 6).

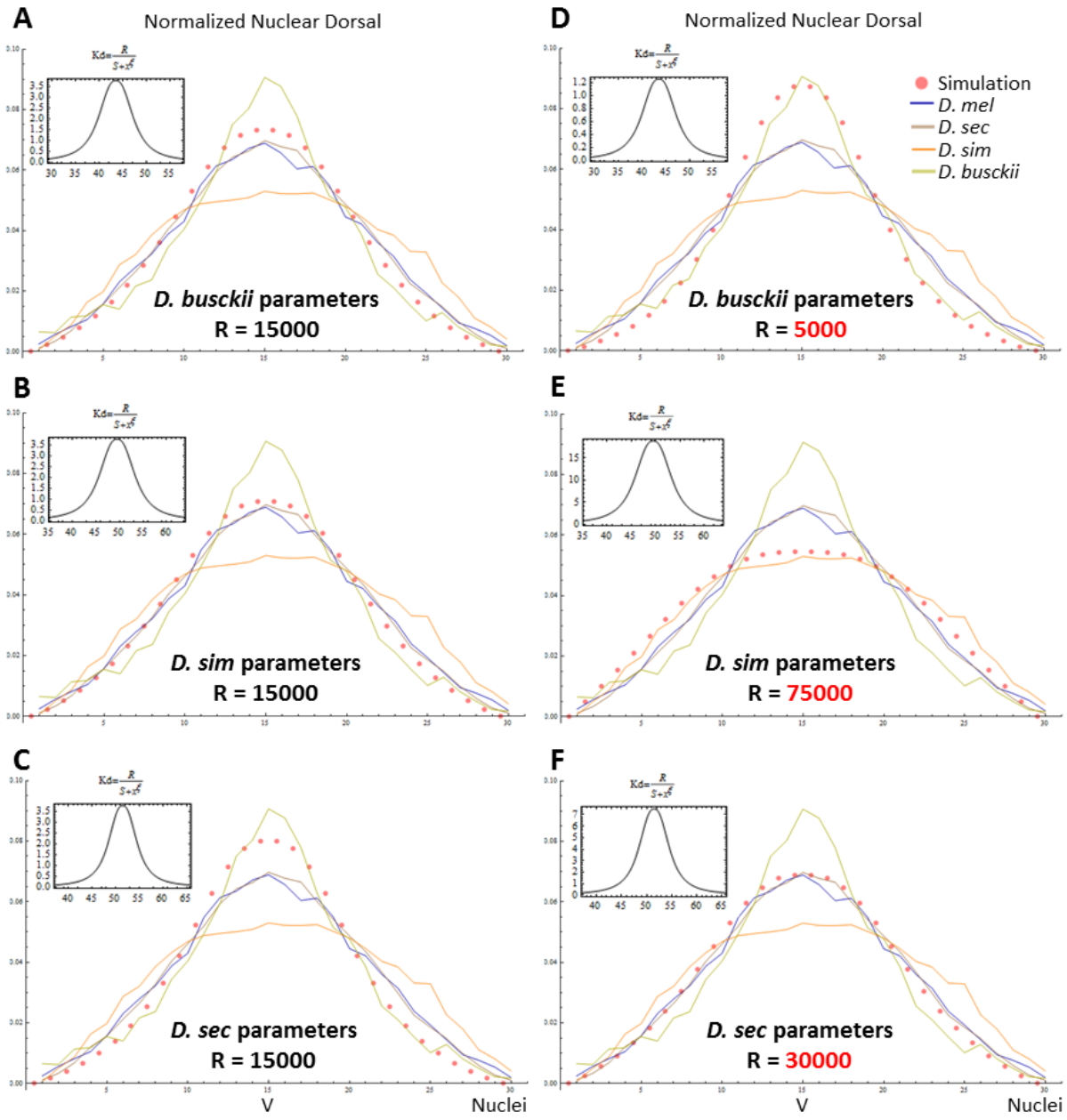
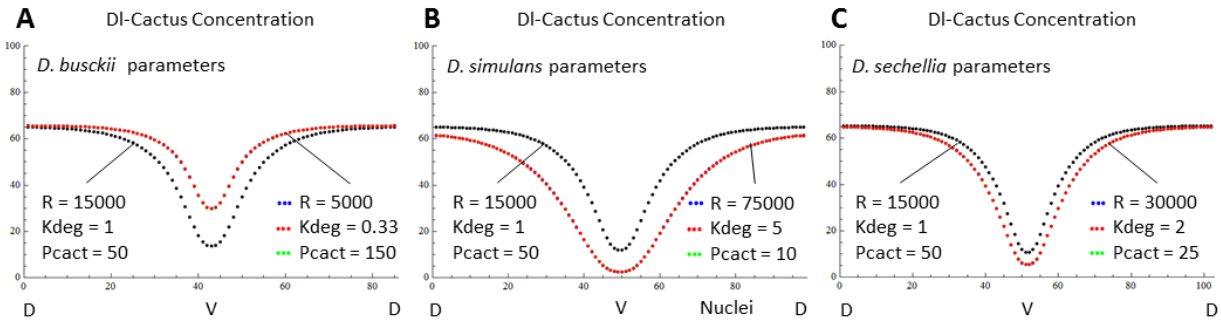


Figure 8. Adjustment of parameters involved in Toll signaling allow the reproduction of species-specific Dl gradients. (A-C) Simulation of *D. busckii* (A), *D. simulans* (B) and *D. sechellia* (C) Dl gradients taking into consideration species-specific embryo size and nuclei density and radius shown in Table 2. (D-F) Simulation of the species gradients considering additional changes in the value of R, a parameter that regulates the amplitude of the Toll signaling gradient. Insets show the Toll signaling gradient (represented by Kd) for the 30 most ventral cells. Note the difference in amplitude of Kd from (A-C) to (D-F). V: ventral midline. Experimental Dl gradients (average) were reproduced from Chahda *et al* [4].



Supplementary Figure 6. Changes in R, Kdeg or Pcact according to the species-specific values shown in Table 2 result in similar distributions of non-normalized Dl-Cactus. Black dotted lines indicate simulations using species-specific nuclear size, density and embryo size, while colored lines indicate simulations with additional changes in either R, Kdeg or Pcact, according to the values shown in the figure and in Table 2. Note that the colored lines are superimposed, showing that the parameter alterations that result in the same Dl nuclear gradient shape also result in the same Dl-Cactus gradient.

We next tested changing the value of R based on experimental data. Previously, we found the existence of species-specific ranges of Toll peak activation, corresponding to the extent of *sna* domain along the entire DV embryo circumference. *D. busckii* has the smallest percent in mesodermal arc-length of 17%, followed by *D. melanogaster* (21%), *D. sechellia* (26%) and *D. simulans* (27%). Thus, we used this experimental data to adjust the parameter R for each species, which controls the amplitude of Kd and the range of the Toll signaling strength extending dorsally from the midline. Under these conditions, the reproduction of species gradients is slightly improved, but does not fully fit the experimental gradient (Figure 9A-C). A good fit between the model and our quantifications was obtained by additionally changing parameters related to Cactus regulation (Figure 9D-F). Either increasing Kdeg or reducing Pcact in both *D. simulans* and *D. sechellia* allowed the correct simulation of their gradients. In contrast, *D. busckii* gradient could be reproduced by the model by either decreasing Kdeg or increasing Cactus production.

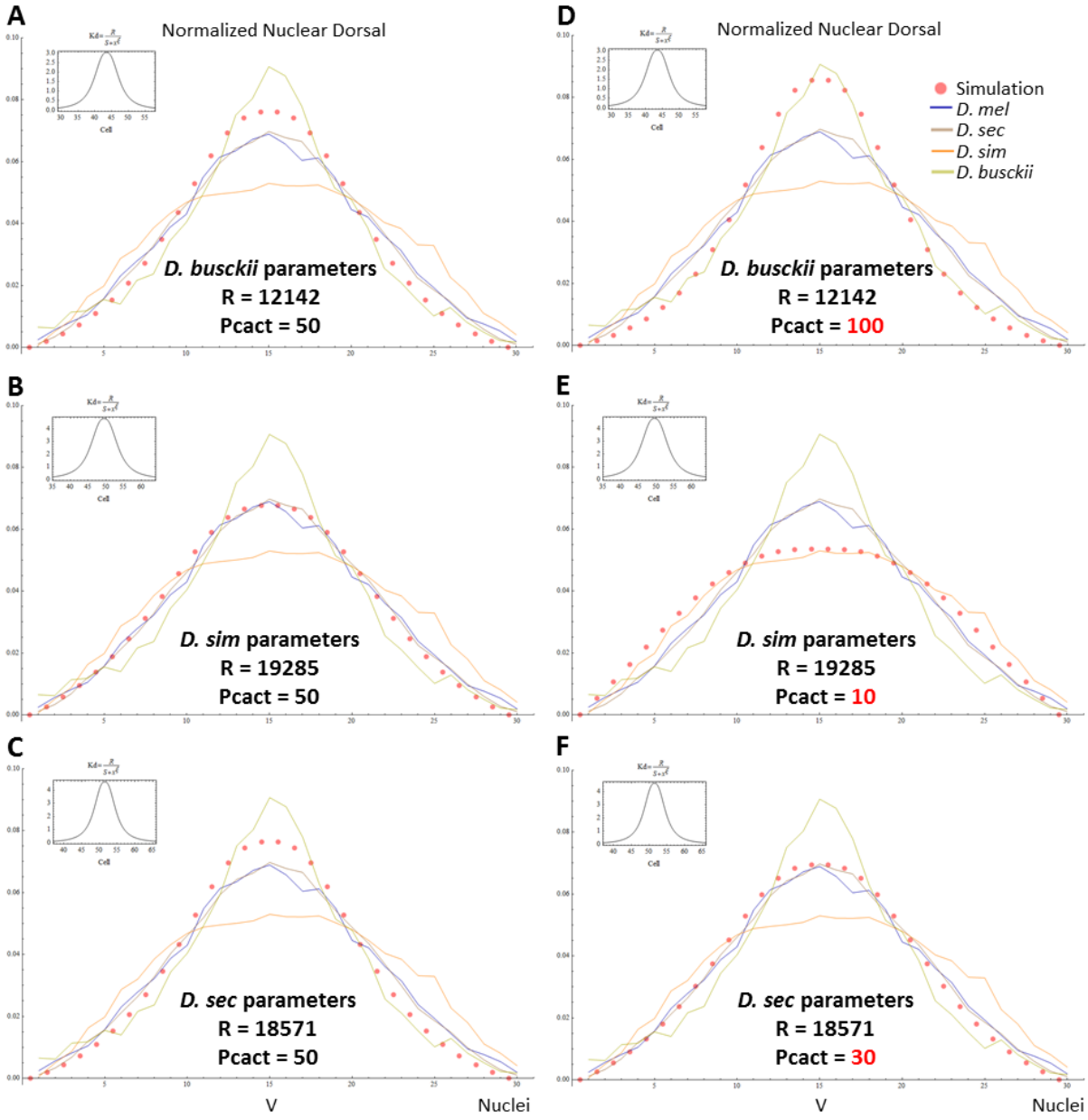


Figure 9. Combination of adjustment of R and Cactus production allow the reproduction of species-specific Dl gradients. (A-C) Simulation of *D. busckii* (A), *D. simulans* (B) and *D. sechellia* (C) Dl gradients taking into consideration species-specific embryo size and nuclei density and radius shown in Table 2 and values of R based on the mesodermal arc-length. (D-F) Simulation of the species gradients considering additional changes in production of Cactus (Pcact). Insets show the Toll signaling gradient (represented by Kd) for the 30 most ventral cells. V: ventral midline. Experimental Dl gradients (average) were reproduced from Chahda *et al* [4].

The results obtained from the use of a computational model allow us to make some conclusions about the behavior of gradient formation and evolution of the Toll signaling pathway in the *Drosophila* species tested. The most related species, *D. simulans* and *D. sechellia*, appear to have diverged in the Toll signaling pathway regulation from their ancestral species, *D. melanogaster.* As suggested in our previous work [4], the range of peak Toll activation became broader in both *D. simulans* and *D. sechellia*, compared to *D. melanogaster.* Our present data suggest that other components of the Toll pathway affecting production of Cactus or its degradation, also diverged in the newest species. Despite the fact that *D. simulans* and *D. sechellia* have particular egg sizes and Dl gradient distributions, the model reveals that these species might share similarly modified components of the Toll pathway inherited from their common ancestor.

Discussion

Scaling of morphogenetic gradients is a significant problem to developmental biology and regeneration fields. During the processes of growth and regeneration of amputated limbs, gene expression patterns regulated by morphogens must adjust to ongoing changes in tissue scale. From an evolutionary standpoint, variation in organismal size also presents a challenge on how gene expression patterns can accommodate to new embryonic sizes without compromising the formation of key cell types and the viability of the organism. In this work, we focused on the scaling of DV patterning in the embryo by analyzing how the shape of the Dl morphogen gradient responds to physical and biochemical changes observed in mutants and different *Drosophila* species.

The Kanodia model can reproduce mutant gradients when Dl diffusion and nuclear export rates are higher than previously assumed

Kanodia *et al* developed a mathematical model of Dl gradient formation using wt *D. melanogaster* experimental data to constrain the model parameter set. Here, we showed that their model and parameter set could correctly simulate the gradients of mutants and other *Drosophila* species after some improvements were made to the model. We removed the non-dimensionalization of the parameters in order to make the model more flexible and allow testing changes of individual parameter values. First, we sought to test our hypothesis that changes in nuclei density and size observed in two *D. melanogaster* mutants, ssm and gyn, are sufficient to generate the mutant Dl gradient shapes. By changing the parameter values of number of nuclei and nuclear radius obtained from the mutants, the adjusted model was able to reproduce the flattened ssm gradient, but not the steep gyn gradient (Figure 3). After removing the non-dimensionalization of the original model, we could test different combinations of parameters other than just nuclei density and size and were able to appropriately reproduce each mutant gradient. To help us validate our new parameter set, we also simulated the Dl gradient of embryos derived from dl^{-}/dl^{+} mothers, which are expected to have initial concentrations of Dl deposited in the embryo reduced by half. By taking this approach, we identified two key parameters involved in the establishment of the Dl gradient shape that had different values than the original model, namely the Dl diffusion and nuclear export rates.

The Kanodia model based the Dl diffusion rates on photobleaching experiments performed with Dl-GFP, which concluded the syncytial blastoderm is highly compartmentalized even in

the absence of cell membranes [15]. It is possible that the large fusion protein Dl-GFP diffuses at much lower rates than endogenous Dl. Therefore, we reasoned that increasing the diffusion rates of Dl between compartments would be a reasonable assumption with possibly a large effect over the Dl gradient formation. By increasing this parameter, we actually improved the fit of our simulations of the wt gradient, and were able to effectively reproduce the gradients of ssm and dl^{-}/dl^{+} embryos. For gyn gradient simulation, we found that increasing Dl diffusion also sharpened the gradient improving its fit, although it was not sufficient to fully reproduce the experimentally obtained gradient. Evidence that Dl molecule may diffuse more than just one compartment away comes from a study that overexpressed an activated form of Toll receptors in the embryo head [16]. In this experimental setting, *snail* is ectopically expressed in the head in addition to its normal ventral stripe. However, 50% of these embryos display a large gap in *sna* ventral stripe, juxtaposed to the *sna* ectopic expression in the head. While Huang *et al* do not genotype individual embryos, we interpret that this 50% class should correspond to embryos that carry two copies of the activated Toll construct. In this case, a local increase of Dl transport to the nucleus in the head could create an influx of endogenous Dl to this site of high Toll activation, causing Dl depletion in the adjacent region. Such effect is expected to be downstream of Toll receptor, since the construct expresses a constitutively activated form of Toll that activates the pathway independently from the ligand. A similar sequestering effect has been seen for Dpp ligand in embryos that overexpress Thickveins (Tkv) receptors in the head, which result in increased PMad (phosphorylated Mad – Mothers against Dpp) at the site of ectopic Tkv expression but a decrease in PMad in the region adjacent to it [17]. The effect over Dl dislocation to the site of ectopic Toll expression seem to encompass 7 cell diameters, therefore, at least 3 times more than a diffusion of Dl occurring only between neighboring compartments.

Another departure from the Kanodia model was to employ a higher Dl nuclear export rate, given that the Dl-GFP construct used to estimate this value has a deletion on a putative export sequence of the Dl protein [18]**.** By simultaneously increasing Dl diffusion rates and Dl nuclear export rates, we significantly improved the gyn gradient simulation with a very good sharpening and a slightly lower peak than the experimental. The combined change in Dl diffusion and nuclear export rate had no effect on the simulations of wt and ssm gradient, except for a small increase of the Dl peak of dl^{-}/dl^{+} mutant gradient. A nearly perfect fit for the gyn gradient was obtained when increasing the radius of the embryo from 90 to 110 µm. While we could not find the best parameter set for gyn and these mutant embryos are not expected to differ in size from wt embryos, this result shows that the model is sensitive to

small changes of embryo geometry below 23%, which is less than the differences in DV axis among the other *Drosophila* species tested.

Other changes in the model had negligible effects. For instance, we tested the effect of cortical layer size in modifying the Dorsal gradient. Kanodia *et al* assumed that the width of the cortical layer was constant throughout development, even though it was previously shown that the width increases as the embryo develops, going from 11 µm at nc11 to 31 µm at nc14 [19]. We also directly measured the width of the cortical layer in the mutants and found it to be larger in ssm and gyn in comparison to wt (data not shown). Hence, we tested if changes in the cortical layer width would improve our simulations. We found that changing the width of the cortical layer does not significantly affect the shape of the Dl gradient (Supplementary Figure 4).

Model predictions about changes in the Toll pathway in *Drosophila* species.

As expected, our model predicts that the Dl gradient is highly sensitive to the shape of the Toll signaling gradient, represented by the parameter ξ used to model the space-dependent reaction rate constant for dissociation of Dl-Cactus (Supplementary Figure 4). Additionally, parameters indirectly involved in the rate of decay of the Toll signal, like the number of nuclei in a DV cross-section and embryo radius also considerably affect the shape of the Dl gradient. Other parameters, like the total number of nuclei in the embryo, the AP length of the embryo and the absolute initial concentrations of Dl, Dl-Cactus or Cactus, have an almost irrelevant influence on the shape of the gradient. These results point to the preponderance of embryo geometry and nuclei arrangement along the axes above overall embryo volume and number of nuclei in regulating the Dl gradient shape. One could argue that this result is an artifact, because our model does not simulate diffusion along the AP axis. Such diffusion would, however, simply affect the absolute concentration of Dl in a given cross-section. Given that the model is insensitive to the initial concentration of Dl, such simulation is unlikely to change the model outcome.

The model makes important predictions about the mechanisms for Dl gradient formation in the species. On a first analysis, it is surprising to find that the gradients of *D. sechellia* and *D. melanogaster* are identical in distribution despite their divergence of 4.5 million years, whereas the gradient of *D. simulans*, a species much more closely related to *D. sechellia*, has acquired a completely different shape after an estimated divergence of only 0.5 million years. The model reconciles this finding by revealing that these quite different gradient shapes can

in fact be generated by a similar dynamics of Toll signaling among *D. simulans* and *D. sechellia*, i.e. a similar range of peak activation and similar Cactus regulation. The model also confirms that other physical changes independent from the Toll pathway also play an important role in the formation of the Dl gradient.

Given the model simplification of the Cactus degradation pathways, it is not possible to define exactly which components might be changed in the species. Future work could test if the formation of the Dl gradient evolves by small additive modifications in many nodes of the Toll signaling pathway, or if mutations in one component at a time leading to changes in the regulation of the pathway (e.g. increased Cactus production) are fixed in the species lineages.

Materials and Methods

Fly stocks and genetic crosses

yw D. melanogaster was used as wild-type. Haploid embryos were generated by mating homozygous females *w, ssm* (gift from James Erickson) [20,21] to wild-type males. 100% of the progeny of this cross develop as haploids. Triploid embryos were generated by mating *gynogenetic-2; gynogenetic-3* (*gyn*) [22] double homozygous females (Bloomington Stock Center) to wild-type males. About 12% of the embryos derived from this cross develop as triploids. To maintain the 12% triploidy rate, *gyn2; gyn3* females were mated to sterile males (*ms*(3)K81[1]) and F1 females that developed parthenogetically [22] were used to generate triploid embryos, as described above.

Measurements of nuclear size

Fixed embryos stained with anti-laminin were mounted longitudinally with glass beads (150-210 µm size, Polysciences), to prevent flattening caused by the coverslip. Confocal slices were taken from the embryo surface to its mid-section (Zeiss LSM700) and nuclei diameter was determined using ImageJ software. The most spherical nuclei from early-blastoderm embryos were analyzed. Due to small sample size, additional measurements were taken from cross-sections of trunk regions of embryos stained with DAPI nuclear dye. Sample sizes for both ssm and gyn are the following: n=58 nuclei from n=4 embryos (n=1 embryo stained with anti-laminin and n=3 embryos stained with DAPI).

Reproduction and modification of the Kanodia model in Mathematica

Nondimensionalized model of nc10-14. The nondimensionalized model was reproduced as described by Kanodia *et al* [8]. To simulate gyn and ssm mutant gradients, we used the same equations (Supplementary Methods), but changed parameter values as described below.

Genotype-specific parameters. While nuclei radius at the last nuclear cycle was directly measured for gyn and ssm mutants, the nuclei density was based on the number of cycles until cellularization. Considering that wt embryos have about 6000 nuclei at nc14, we assumed that gyn embryos have 3000 nuclei at nc13 and that ssm embryos have 12000 nuclei at nc15. The latter is supported by Edgar *et al* [23]. To determine the number of nuclei in a cross-section, we used the same method employed by Kanodia: after each division, the

number of nuclei in a cross-section is multiplied by $\sqrt{2}$, being 100 the number of nuclei in a wt cross-section at nc14 [8].

Besides nuclei density and radius at the last nuclear cycle, other parameters are likely to change between genotypes. One of them is developmental timing. Haploid embryos were shown to have similar developmental timing to wt until nc12, after which the cycles get shorter. The haploid nc13/14 lasts about as long as the wt nc12/13 [23,24]. Limited information is available for triploid embryos, so we have assumed that gyn development is wt until nc13, and that nc13 spans the duration of wt nc13 and 14. Specific values used in the model are shown in Supplementary Table 4.

The nuclear radius prior to the last cycle is also unknown for the mutants. This measurement is challenging to be done even in wt embryos, because each cycle lasts about 10 minutes and nuclear size changes substantially from early-interphase to late-anaphase. However, the haploid nucleus seems to be similar to wt until nc14 [24,25], suggesting that the extra division is responsible for nuclear reduction. The fact that the triploid nucleus is larger than wt at nc13 suggests that the triploid nucleus is generally larger than wt. Hence, the following assumptions were made: (1) the nuclear radius of ssm embryos is wt until its extra division, after which it adopts the value determined experimentally (2.3 μm); (2) gyn nuclear radius is given by an interpolation of wt values adjusted to the ratio gyn/wt nuclear radius at nc13 (5.45/4.2). A graphical representation of nuclear radius versus time for all genotypes can be seen in Supplementary Figure 7.

Dimensionalized model of the last nuclear cycle. Although the nondimensionalization reduces the number of parameters, it also hinders the simulation of embryos with a different geometry from wt. Since we were interested in simulating these differences, we decided to use dimensionalized equations, which are the original mass-balance equations written by Kanodia (Supplementary Table 1). Additionally, to better represent the changes in embryo volume between species, we have decided to keep the cross-section as a circle, instead of linearizing it. Hence, instead of a row of rectangular cell compartments with no-flux boundary conditions, our simulated cross-section is made of circular trapezoids (or thick annular sectors) organized in a circle (Supplementary Figure 8). Details of the modifications are described in Supplementary Methods.

Supplementary Materials and Methods

Kanodia model description

Mass-balance equations.

Supplementary Table 1. Term-by-term description of the model differential equations.

$\frac{d(V_n C^h_{Dl,n})}{dt} =$	$+A_n k_i C^h_{Dl,c}$		$-A_n k_e C^h_{Dl,n}$		
Time rate of change of the amount of Dl in the nucleus of the compartment h	Transport of Dl from the cytoplasm to the nucleus		Transport of Dl from the nucleus to the cytoplasm		
$\frac{d(V_c C^h_{Dl,c})}{dt} =$	$+\Gamma A_m(C^{h+1}_{Dl,c} - 2C^h_{Dl,c} + C^{h-1}_{Dl,c})$	$+k_D C^h_{Dl-cact,c} V_c$	$-k_b C^h_{Dl,c} C^h_{cact,c} V_c$	$-\mathrm{A_n k_i C^h_{Dl,c}}$	$+\mathrm{A_n k_e C^h_{Dl,n}}$
Time rate of change of the amount of Dl in the cytoplasm of the compartment h	Transport of Dl between the cytoplasm of the compartment h and the adjacent compartments	Dissociation of the Dl-Cact complex	Association of Dl and Cactus to form the Dl-Cact complex	Transport of Dl from the cytoplasm to the nucleus	Transport of Dl from the nucleus to the cytoplasm
$\frac{d(V_c C^h_{Dl-cact,c})}{dt} =$	$+\Gamma A_m(C^{h+1}_{Dl-cact,c} - 2C^h_{Dl-cact,c} + C^{h-1}_{Dl-cact,c})$			$-k_D C^h_{Dl-cact,c} V_c$	$+k_b C^h_{Dl,c} C^h_{cact,c} V_c$
Time rate of change of the amount of Dl-Cact complex in the cytoplasm of the compartment h	Transport of Dl-Cact complex between the cytoplasm of the compartment h and the adjacent compartments			Dissociation of the Dl-Cact complex	Association of Dl and Cactus to form the Dl-Cact complex
$\frac{d(V_c C^h_{cact,c})}{dt} =$	$+\Gamma A_m(C^{h+1}_{cact,c} - 2C^h_{cact,c} + C^{h-1}_{cact,c})$	$+k_D C^h_{Dl-cact,c} V_c$	$-k_b C^h_{Dl,c} C^h_{cact,c} V_c$	$+P_{cact} V_c$	$-k_{Deg} C^h_{cact,c} V_c$
Time rate of change of the amount of free Cactus in the cytoplasm of the compartment h	Transport of free Cactus between the cytoplasm of the compartment h and the adjacent compartments	Dissociation of the Dl-Cact complex	Association of Dl and Cactus to form the Dl-Cact complex	Production of Cactus	Degradation of Cactus
k_D =	$R / (S + x^{\xi})$				
Space-dependent reaction rate constant for dissociation of the Dl-Cactus complex, representing the Toll signaling gradient	R and S determine the maximum value of k_D, i.e., the amplitude of the Toll signaling gradient, while ξ represents the rate of decay of k_D with an increase in x (the distance from the ventral midline along the DV axis).				

Supplementary Table 2. Description of the model's variables and parameters.

Model Variables	
x	Distance from the ventral midline
$C^h_{Dl,n}$	Concentration of Dl in the nucleus of compartment h
$C^h_{Dl,c}$	Concentration of Dl in the cytoplasm of compartment h
$C^h_{Dl-cact,c}$	Concentration of the Dl-Cactus complex in the cytoplasm of compartment h
$C^h_{cact,c}$	Concentration of free Cactus in the cytoplasm of compartment h

Model Parameters	
k_i	Rate at which Dl enter the nucleus
k_e	Rate at which Dl exits the nucleus
Γ	Transport rate of Dl, Dl-Cactus or Cactus between adjacent compartments
k_b	Association rate of Dl and Cactus to form the Dl-Cactus complex
P_{cact}	Rate of Cactus production
k_{Deg}	Rate of Cactus degradation
k_D	Dissociation rate of the Dl-Cactus complex – the value of this parameter is space dependent as it represents the Toll signaling gradient
R	
S	Parameters that describe the space-dependence of k_D
ξ	

Time-dependent parameters	
V_c	Volume of the cell compartment cytoplasm
V_n	Volume of the cell compartment nucleus
A_n	Surface area of the nucleus
A_m	Surface area between two adjacent compartments

Supplementary Table 3. Parameters and equations used for the model's nondimensionalization.

Scaling factors	
L	Length of the embryo from the ventral to the dorsal midline
T	Total developmental timing from the beginning of cycle 10 to the end of cycle 14
$C^o_{Dl-catc}$	Concentration of Dl-Cactus complex at the beginning of cycle 10
C^o_{catc}	Concentration of free Cactus at the beginning of cycle 10
A^{14}_n	Area of the nucleus at the end cycle 14
A^{14}_m	Surface area between two adjacent compartments at cycle 14
V^{14}_n	Volume of nucleus at the end cycle 14

Dimensionless variables and time-dependent parameters			
$\tau = \frac{t}{T}$	$\bar{C}^h_{Dl,n} = \frac{C^h_{Dl,n}}{C^o_{Dl-cact}}$	$\bar{C}^h_{Dl-cact,c} = \frac{C^h_{Dl-cact,c}}{C^o_{Dl-cact}}$	
$z = \frac{x}{L}$	$\bar{C}^h_{Dl,c} = \frac{C^h_{Dl,c}}{C^o_{Dl-cact}}$	$\bar{C}^h_{cact,c} = \frac{C^h_{cact,c}}{C^o_{cact}}$	
$\bar{V}_n = \frac{V_n}{V^{14}_n}$	$\bar{V}_c = \frac{V_c}{V^{14}_n}$	$\bar{A}_n = \frac{A_n}{A^{14}_n}$	$\bar{A}_m = \frac{A_m}{A^{14}_m}$

Dimensionless parameters	
$\sigma = \frac{A^{14}_n k_i}{V^{14}_n} T$	Nuclear import
$\mu = \frac{A^{14}_n k_e}{V^{14}_n} T$	Nuclear export
$\lambda = \frac{A^{14}_m \Gamma}{V^{14}_n} T$	Transport across compartments
$\gamma = k_b C^o_{cact} T$	Association of Dl-Cactus complex
$\alpha = k_{Deg} T$	Degradation of Cactus
$\psi = \frac{C^o_{Dl-cact}}{C^o_{cact}}$	Relative levels of Dl and Cactus
$\beta = \frac{RT}{L^\xi}$	
$\varphi = \frac{S}{L^\xi}$	Characterization of the spatial pattern for Dl-Cactus dissociation rate
ξ	

Zero-flux boundary conditions. Kanodia simulates only half-circunference of the embryo and uses zero-flux boundary conditions at the most ventral and most dorsal cells. In other words, there is no diffusion beyond these cells and they have only one neighboring cell compartment, with which their exchange rates are doubled. Hence, for all times:

At x=0 (ventral midline), $C_{sp,c}^{h=1} = C_{sp,c}^{h=2}$

At x=L/2 (dorsal midline), $C_{sp,c}^{h=n+1} = C_{sp,c}^{h=n}$

sp ∈ [Dl, Dl-Cactus, Cactus]

Initial conditions. For all x ∈ [0, L]

For nc 10: $C_{catc,c}^{h} = C_{catc,c}^{0} = P_{cact} / k_{Deg}$ $\quad C_{Dl-catc,c}^{h} = C_{Dl-catc,c}^{o}$ $\quad C_{Dl,c}^{h} = C_{Dl,n}^{h} = 0$

For nc 11-14: based on concentration profiles established at the end of previous cycle (see [8] for details).

Nondimensionalized equations. As it can be seen from Supplementary Table 1, the kinetic relationships represented by the differential equations are accompanied by several unknown rate constants and properties related to the shape of the cellular compartments (summarized in Supplementary Table 2). Some of these unknown values were combined into 9 dimensionless parameters (Supplementary Table 3), while others were estimated with time-dependent equations. The fully nondimensionalized equations are shown below. All equations and tables are based on the supplementary information provided by Kanodia *et al* [8].

$$\frac{d\left(\overline{V_n}\,\overline{C}_{\mathrm{dl},n}^{h}\right)}{d\tau} = \sigma\overline{A_n}\,\overline{C}_{\mathrm{dl},c}^{h} - \mu\overline{A_n}\,\overline{C}_{\mathrm{dl},c}^{h}$$

$$\frac{d\left(\overline{V_C}\,\overline{C}_{\mathrm{dl},c}^{h}\right)}{d\tau} = \lambda\overline{A_m}\left(\overline{C}_{\mathrm{dl},c}^{h+1} - 2\overline{C}_{\mathrm{dl},c}^{h} + \overline{C}_{\mathrm{dl},c}^{h-1}\right) + \psi \mathrm{k}_D T\overline{V_C}\,\overline{C}_{\mathrm{dl-cact},c}^{h} - \gamma\psi\overline{V_C}\,\overline{C}_{\mathrm{dl},c}^{h}\,\overline{C}_{\mathrm{cact},c}^{h} - \sigma\overline{A_n}\,\overline{C}_{\mathrm{dl},c}^{h} + \mu\overline{A_n}\,\overline{C}_{\mathrm{dl},c}^{h}$$

$$\frac{d\left(\overline{V_C}\,\overline{C}_{\mathrm{dl-cact},c}^{h}\right)}{d\tau} = \lambda\overline{A_m}\left(\overline{C}_{\mathrm{dl-cact},c}^{h+1} - 2\overline{C}_{\mathrm{dl-cact},c}^{h} + \overline{C}_{\mathrm{dl-cact},c}^{h-1}\right) - \psi \mathrm{k}_D T\overline{V_C}\,\overline{C}_{\mathrm{dl-cact},c}^{h} + \gamma\psi\overline{V_C}\,\overline{C}_{\mathrm{dl},c}^{h}\,\overline{C}_{\mathrm{cact},c}^{h}$$

$$\frac{d\left(\overline{V_C}\,\overline{C}_{\mathrm{cact},c}^{h}\right)}{d\tau} = \lambda\overline{A_m}\left(\overline{C}_{\mathrm{cact},c}^{h+1} - 2\overline{C}_{\mathrm{cact},c}^{h} + \overline{C}_{\mathrm{cact},c}^{h-1}\right) + \psi \mathrm{k}_D T\overline{V_C}\,\overline{C}_{\mathrm{dl-cact},c}^{h} - \gamma\psi\overline{V_C}\,\overline{C}_{\mathrm{dl},c}^{h}\,\overline{C}_{\mathrm{cact},c}^{h} + \alpha\overline{V_C} - \alpha\overline{V_C}\,\overline{C}_{\mathrm{cact},c}^{h}$$

$$k_D T = \frac{\beta}{\phi + z^{\xi}}$$

Genotype-specific parameter values. To determine the volume of the cell compartments, Kanodia approximated the cortical region of the embryo as the space between two ellipsoids. He then divided the volume of the cortical region by 6000 cell compartments to estimate the volume of an individual compartment at nc14. The final value obtained by Kanodia was 865 μm^3. This value doubles for each previous nuclear cycle, because the number of

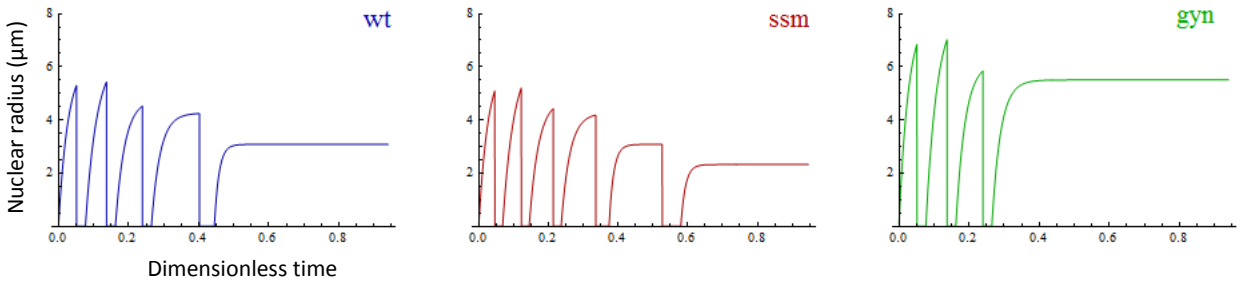
compartments halves. The nuclear volume, on the other hand, was determined by a time-dependent function based on the measurements by Gregor *et al* [19]. The volume of the cytoplasm was calculated by subtracting the nuclear volume from the compartment volume. Additionally, the surface area between two adjacent compartments is divided by $\sqrt{2}$ at the end of each nuclear cycle, because the width of the cortical layer is assumed to be constant by Kanodia *et al*.

To simulate ssm and gyn using the nondimensionalized equations, we assumed that all genotypes had the same embryo size. Hence, the cell compartments of ssm at nc15 had half the volume of wt cell compartments at nc14, while gyn compartments at nc13 had the same volume as wt cell compartments at nc13. Also, developmental timing and nuclear radius were adjusted for ssm and gyn according to Supplementary Table 4 and Supplementary Figure 8.

Supplementary Table 4. Genotype-specific parameter values.

Genotype	Last cycle before gastrulation	Number of nuclei at last cycle*	Nuclear radius at nc14 (µm)	Total duration of nuclear cycles (min)					
				10	11	12	13	14	15
wt	14	100	3.05	9	10	12	21	65	-
ssm	15	142 (100 * $\sqrt{2}$)	2.3	9	10	12	18	27	55
gyn	13	72 (100 / $\sqrt{2}$)	5.45	9	10	12	86	-	-

*In a full circumference of a DV cross-section. At each nc, this number is multiplied by $\sqrt{2}$ [8].

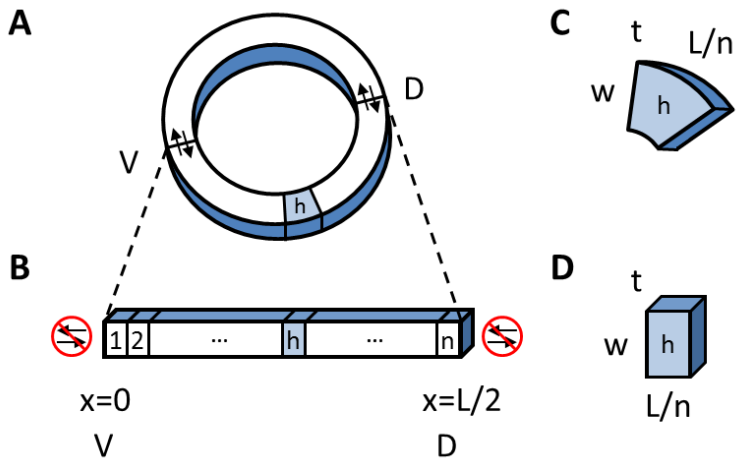


Supplementary Figure 7. Time-dependent nuclear radius dynamics for wt (blue), ssm (red) and gyn (green).

Modified model description

In the original model, only a DV semi-circumference was modeled, which was approximated as a rectangle with no-flux boundary conditions. In combination with the assumption that the cortical layer width is constant throughout development, this simplification allowed Kanodia *et al* to easily adjust the values of Am and Vc from nc 10 to 14. However, it is known that the cortical layer width is not constant throughout development [19], and that it is altered in gyn and ssm mutants, as well as in different species (Chahda et al, unpublished data). As the width of the cortical layer increases, the error resulting from the linearization of the cross-section also increases. Hence, we developed a modified model in which a full DV cross-section was simulated, and the curved shape of the cell compartments was preserved (Supplementary Figure 9).

Additionally, to maximize the simulation flexibility, the modified model simulates a single nuclear cycle, and all parameters shown in Table 1 can be manually adjusted. The volume of each cell compartment is calculated based on the size of the embryo, the width of the cortical layer and the total number of nuclei in the embryo, similarly to the method employed by Kanodia. Differently from the Kanodia model, the initial conditions are not adjusted according to previous cycles, because previous cycles are not simulated. Our decision to make this modification was based on our findings that early cycles do not affect the shape of the Dl gradient at nc14 (Supplementary Figure 3), and it significantly reduces the computational requirements of the model. Also, we did not simulate time-dependent changes in nuclear radius. Instead, the nuclear radius is set to a constant value. All simulations after Supplementary Figure 4 and Figure 5 were done using the modified model.



Supplementary Figure 8. Comparison between the original and the modified model. (A) Cross-section scheme representing normal flux between its two halves (arrows) and a single cell compartment named h. (B) Linearized half cross-section, with compartments 1 (ventral most cell) to n (dorsal most cell), and no-flux boundary conditions (crossed arrows). Note that L/2 was used to represent the distance from the ventral midline to the dorsal midline in this figure, but Kanodia et al represents the same distance with L. (C) A single compartment according to the modified model. (D) A single compartment according to the original Kanodia model. w: width of the cortical layer; t: thickness of the cortical layer; n: number of compartments in a full DV cross-section; L: length of the DV axis.

Acknowledgements

We thank for funds received from National Science Foundation (IOS 1051662) to CMM and from the Brazil Science Without Borders scholarship (sponsored by CAPES) to PA. We also thank Vera Lúcia da Silva Valente for critical reading of the manuscript.

References

1. Doe CQ (1992) Molecular markers for identified neuroblasts and ganglion mother cells in the *Drosophila* central nervous system. Development 116: 855–863.

2. Thomas JB, Bastiani MJ, Bate M, Goodman CS (1984) From grasshopper to *Drosophila*: a common plan for neuronal development. Nature 310: 203–207.

3. Whitington PM (1996) Evolution of neural development in the arthropods. Seminars in Cell & Deveopmental Biology 7: 605–614. doi:http://dx.doi.org/10.1006/scdb.1996.0074.

4. Chahda JS, Sousa-Neves R, Mizutani CM (2013) Variation in the dorsal gradient distribution is a source for modified scaling of germ layers in *Drosophila*. Current biology : CB 23: 710–716. doi:10.1016/j.cub.2013.03.031.

5. Belu M, Mizutani CM (2011) Variation in mesoderm specification across Drosophilids is compensated by different rates of myoblast fusion during body wall musculature development. PloS one 6: e28970. doi:10.1371/journal.pone.0028970.

6. Gregor T, Bialek W, de Ruyter van Steveninck RR, Tank DW, Wieschaus EF (2005) Diffusion and scaling during early embryonic pattern formation. Proc Natl Acad Sci U S A 102: 18403–18407.

7. Schlenke T, Begun D (2003) Natural selection drives *Drosophila* immune system evolution. Genetics 164: 1471–1480.

8. Kanodia JS, Rikhy R, Kim Y, Lund VK, DeLotto R, et al. (2009) Dynamics of the Dorsal morphogen gradient. Proceedings of the National Academy of Sciences of the United States of America 106: 21707–21712. doi:10.1073/pnas.0912395106.

9. Foe VE, Alberts BM (1983) Studies of nuclear and cytoplasmic behaviour during the five mitotic cycles that precede gastrulation in *Drosophila* embryogenesis. Journal of cell science 61: 31–70.

10. Drier E, Govind S, Steward R (2000) Cactus-independent regulation of Dorsal nuclear import by the ventral signal. Current Biology 10: 23–26.

11. Belvin MP, Jin Y, Anderson K V (1995) Cactus protein degradation mediates *Drosophila* dorsal-ventral signaling. Genes & Development 9: 783–793. doi:10.1101/gad.9.7.783.

12. Whalen a M, Steward R (1993) Dissociation of the dorsal-cactus complex and phosphorylation of the dorsal protein correlate with the nuclear localization of dorsal. The Journal of cell biology 123: 523–534.

13. Markow TA, Beall S, Matzkin LM (2009) Egg size, embryonic development time and ovoviviparity in *Drosophila* species. J Evol Biol 22: 430–434.

14. Miles C, Lott S, Hendriks CL (2011) Artificial selection on egg size perturbs early pattern formation in *Drosophila melanogaster*. Evolution; international journal of organic evolution 65: 33–42. doi:10.1111/j.1558-5646.2010.01088.x.Artificial.

15. DeLotto R, DeLotto Y, Steward R, Lippincott-Schwartz J (2007) Nucleocytoplasmic shuttling mediates the dynamic maintenance of nuclear Dorsal levels during *Drosophila* embryogenesis. Development 134: 4233–4241.

16. Huang AM, Rusch J, Levine M (1997) An anteroposterior Dorsal gradient in the *Drosophila* embryo. Genes Dev 11: 1963–1973.

17. Mizutani CM, Nie Q, Wan FYM, Zhang Y-T, Vilmos P, et al. (2005) Formation of the BMP activity gradient in the *Drosophila* embryo. Developmental cell 8: 915–924. doi:10.1016/j.devcel.2005.04.009.

18. Xylourgidis N, Roth P, Sabri N, Tsarouhas V, Samakovlis C (2006) The nucleoporin Nup214 sequesters CRM1 at the nuclear rim and modulates NFkappaB activation in *Drosophila*. Journal of cell science 119: 4409–4419. doi:10.1242/jcs.03201.

19. Gregor T, Wieschaus EF, McGregor AP, Bialek W, Tank DW (2007) Stability and nuclear dynamics of the bicoid morphogen gradient. Cell 130: 141–152.

20. Loppin B, Couble P, Bernard C (2000) The Maternal Effect Mutation sesame Affects the Formation of the Male Pronucleus in *Drosophila melanogaster*. Developmental Biology 222: 392–404. doi:10.1006/dbio.2000.9718.

21. Erickson JW, Quintero JJ (2007) Indirect effects of ploidy suggest X chromosome dose, not the X:A ratio, signals sex in Drosophila. PLoS Biol 5: e332.

22. Fuyama Y (1986) Genetics of parthenogenesis in *Drosophila melanogaster*. 11. Characterization of a gynogenetically reproducing strain. Genetics 114: 495–509.

23. Edgar B a, Kiehle CP, Schubiger G (1986) Cell cycle control by the nucleo-cytoplasmic ratio in early *Drosophila* development. Cell 44: 365–372.

24. Lu X, Li JM, Elemento O, Tavazoie S, Wieschaus EF (2009) Coupling of zygotic transcription to mitotic control at the *Drosophila* mid-blastula transition. Development (Cambridge, England) 136: 2101–2110. doi:10.1242/dev.034421.

25. Grosshans J, Müller HAJ, Wieschaus E (2003) Control of cleavage cycles in *Drosophila* embryos by frühstart. Developmental cell 5: 285–294.

Conclusões e Perspectivas

Neste trabalho, reproduzimos e modificamos um modelo matemático da formação do gradiente de Dorsal. Nossas alterações permitiram a reprodução de gradientes mutantes e de outras espécies, além do gradiente selvagem de *D. melanogaster*. Nossas modificações foram baseadas em dados da literatura e em dados experimentais próprios e nos permitiram fazer predições a respeito dos principais fatores responsáveis pela determinação da forma do gradiente de Dorsal. Uma de nossas predições diz respeito à amplitude do gradiente de Toll em espécies-irmãs de *Drosophila*, o que está de acordo com dados experimentais prévios e pode ser testado experimentalmente de forma mais direta. Outra de nossas predições é que a geometria do embrião é mais importante do que o volume para a determinação da forma do gradiente de Dorsal, o que também pode ser testado experimentalmente.

Além disso, nossos resultados indicam certas relações quantitativas entre parâmetros, sugerindo que mudanças em um ou mais parâmetros podem ser compensadas por outros parâmetros. Por exemplo, alterações na cinética de produção e degradação de Cactus podem ser compensadas por alterações na amplitude do gradiente de sinalização Toll. Nossos achados em relação às espécies *D. simulans*, *D. sechellia* e *D. busckii* sugerem que, na evolução dessas espécies, mecanismos distintos atuaram para estabelecer os gradientes de Dorsal de cada espécie. Especificamente, *D. simulans* e *D. sechellia*, espécies que divergiram há apenas 0,5 milhões de anos, parecem ter taxas de produção de Cactus inferiores às de *D. melanogaster*, enquanto *D. busckii*, uma espécie ainda mais divergente, parece ter taxas superiores de produção de Cactus. Tais resultados apontam para abordagens experimentais futuras, como a análise do gene que codifica Cactus dessas espécies, buscando identificar fatores que indiquem um maior ou menor *turnover* dessa proteína.

Adicionalmente, a formação do gradiente de Dorsal em outras espécies ancestrais de *Drosophila*, como *D. yakuba* e *D. santomea*, pode ser futuramente simulada pelo modelo aqui apresentado, visando identificar parâmetros e mecanismos candidatos que sejam importantes para a evolução do estabelecimento das camadas germinativas e para o escalonamento do gradiente de Dorsal.

Portanto, as perspectivas desse trabalho incluem (1) quantificações detalhadas dos parâmetros envolvidos no modelo, (2) validação experimental das predições apresentadas, e (3) simulação do gradiente de espécies ancestrais de *Drosophila*. Extensões do modelo incluindo fatores *upstream* e

downstream da sinalização Toll também podem ser feitas, buscando identificar com maior precisão os mecanismos divergentes entre as espécies. Além disso, a formação de outros gradientes morfogenéticos e limiares para expressão dos genes *vnd*, *ind* e *msh* também podem ser simulados, buscando identificar mecanismos capazes de reproduzir a robustez dos subdomínios do neuroectoderma.