What is meant by "gap genes" and the "co-linearity phenomenon".

If the normal phenotype were the alphabet as shown below:

ABCDEFGHIJKLMNOPQRSTUVWXYZ

Then the phenotype of mutants of the Bicaudal and Bicoid genes would be:

ZYXWUTSRQPONNOPQRSTUVWXYZ

Which had <u>also been seen to develop when strings were tied tightly around the middles of dragon fly early embryos</u>.
(Two mirror-image rear halves, with no anterior organs)

And the phenotypes of **gap gene mutants** would be analogous to these:

ABCDEFGHIJKTUVWXYZ ("lmnopqrs" missing)

ABCDJKLMNOPQRSTUVWXYZ ("efghi" missing)

AABBCCDDEEFFTTUUVVWWXXYYZZ ("ghijklmnopqrs" missing)

Whereas the phenotypes of **pair-rule genes** mutants would be analogous to these

AABBEEFFIIJJMMNNQQRRUUVVYYZZ

CCDDGGHHIILLMMPPQQTTUUXXYYZ

& my best effort to approximate the effect of **segment polarity gene** mutant phenotypes of is the following: (using lower case to suggest small mirror images)

AaCcEeGgIiKkMmOoQqSsUuWwYyZ

Next is my attempt to diagram locations where different Hox genes are transcribed:
(What is meant by "**The phenomenon of co-linearity**".

$H_1\ H_1\ H_1\ H_1\ H_1\ H_1\ H_1\ H_1\ H_1\ H_1\ H_1\ H_1\ H_1\ H_1\ H_1\ H_1\ H_1\ H_1\ H_1\ H_1$
$H_2\ H_2\ H_2\ H_2\ H_2\ H_2\ H_2\ H_2\ H_2\ H_2\ H_2\ H_2\ H_2\ H_2\ H_2\ H_2\ H_2$
$H_3\ H_3\ H_3\ H_3\ H_3\ H_3\ H_3\ H_3\ H_3\ H_3\ H_3\ H_3\ H_3\ H_3\ H_3\ H_3\ H_3$
$H_4\ H_4\ H_4\ H_4\ H_4\ H_4\ H_4\ H_4\ H_4\ H_4\ H_4\ H_4$
$H_5\ H_5\ H_5\ H_5\ H_5\ H_5\ H_5\ H_5\ H_5\ H_5\ H_5\ H_5\ H_5$
* $H_6\ H_6\ H_6\ H_6\ H_6\ H_6\ H_6\ H_6\ H_6\ H_6\ H_6\ H_6\ H_6\ H_6\ H_6\ H_6\ H_6$
$H_7\ H_7\ H_7\ H_7\ H_7\ H_7\ H_7\ H_7\ H_7\ H_7\ H_7$
$H_8\ H_8\ H_8\ H_8\ H_8\ H_8\ H_8\ H_8\ H_8\ H_8\ H_8$
$H_9\ H_9\ H_9\ H_9\ H_9\ H_9\ H_9\ H_9\ H_9$
$H_{10}\ H_{10}\ H_{10}\ H_{10}\ H_{10}$
** $H_{12}\ H_{12}\ H_{12}\ H_{12}\ H_{12}\ H_{12}$
$H_{13}\ H_{13}\ H_{13}\ H_{13}$

* Notice that Hox gene #6 is transcribed in locations more anterior than hox gene #5,
**Also notice that Hox gene #11 doesn't exist.

These correspond to some of the observed oddities of Hox gene expression patterns.

The phenomenon of colinearity:

The locations of hox genes on their chromosome would be

H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9 H_{11} H_{12} H_{13}

Some further complication:

First that the Drosophila hox genes are in two separate bunches on the third chromosome
Which can be diagramed like this:
genegene- H_1 H_2 H_3 H_4 H_5-genegenegene-H_6 H_7 H_8 H_9 H_{11} H_{12} H_{13}-genegenegene

Second, that mammals have four different sets of 9-11 hox genes (per set)
with each of the four sets each being on a different chromosome, sort of like this:

A1, A2, A3, A4, A5, A6, A7, A8, A9, A10, A11, A12, A13
B1, B2, B3, B4, B5, B6, B7, B8, B9, B10, B11, B12, B13
C1, C2, C3, C4, C5, C6, C7, C8, C9, C10, C11, C12, C13
D1, D2, D3, D4, D5, D6, D7, D8, D9, D10, D11, D12, D13

Except that 14 of them don't exist; so the real pattern is the following in mammals

A1, A2, A3, A4, A5, A6, A7, * , A9, A10, A11, * , A13 two lacking
B1, B2, B3, B4, B5, B6, B7, B8, B9, * , * , * , * four lacking
* , * , * , C4, C5, C6, * , C8, C9, C10, C11, C12, C13 four lacking
D1, * , D3, D4, * , * , * , D8, D9, D10, D11, D12, D13 four lacking

The numbering is based on similarity of base sequences, so that saying that
D2, D5, D6 and D7 "don't exist" means that in the D bunch of hox genes, none are as closely homologous to A2 or B2 at those two are to each other, and similarly none in the D group have sequences as similar to A5, B5 and C5 as those three genes are to each other, etc.

What has still **not** been discovered…

A) What mechanism turns on transcription of these genes, at the proper locations and times in developing embryos?
B) Whether, or how, the above mechanism for spatial control of transcription uses or depends on the relative locations of the hox genes on their chromosomes?
C) How hox gene expression controls the locations of organs and tissues in the body.

Sometimes the anterior boundaries of expression of specific hox genes corresponds to particular anatomical boundaries. Examples include boundaries between fingers and toes, which correspond to anterior boundaries of expression of hox genes D13 (between little finger and ring finger, and also between the little toe and the toe next to it), hox gene D12 (between ring finger and middle finger, and also between the toes equivalent to them.), and so on for the anterior-most places where D11 and D10 are expressed, which match the locations between middle finger and index finger, and between index finger and thumb (and equivalent toes)!

In situ staining with labeled single stranded nucleic acids have proven the locations of these and other boundaries of expression of specific hox genes. What is lacking are mechanisms to find out (A above) what mechanism causes hox gene D13 to be transcribed in the tissue that becomes the little finger and little toe, but not to be transcribed in the tissues anterior to that, whereas hox genes D12 is transcribed in the little finger and in the ring finger, and D10 is transcribed in the middle finger, ring finger and also the little finger.

Nor has anybody been able to figure out (and prove) how expression of hox 13 causes the formation of the little finger. (question B above)

The best researcher on such subjects is Prof. **Ann Burke**, who was a member of this department for several years, until she moved to another university. One of her methods used specially engineered viruses into which certain hox genes had been inserted. By opening chicken eggs, and putting these viruses at chosen locations, she was able to change what organs would develop at those locations. Imagine causing a mammal embryo to form an extra little toe at any chosen location! She did experiments logically equivalent to that. Prof. Burke's PhD student, **Julie Nowicki** was a TA in this course for two years, and made major discoveries about control of anatomical patterns by hox gene transcription. She also collaborated with Prof. **Alan Feduccia** to test whether the 4 toes of bird feet correspond to toes 1, 2, 3, & 4, as opposed to toes 2, 3, 4 & 5, based on which D hox genes are transcribed in embryonic ostrich (!) feet. This was (very) interesting as a test of hypotheses about which particular group of reptiles birds evolved from (whether from dinosaurs, or not). Their evidence proves "not", but opposition continues.

Some very good scientists (e.g. Slack) believe that co-linearity is not important.

Others think co-linearity must be fundamental to gene control and pattern formation.

A few think they have figured out why it makes functional sense.

Two explanations I have heard proposed: (neither one very convincing)

That there is some kind of concentration gradient down the length of the animal, and it stimulates hox genes' transcription by means of some long range enhancer, the range of which is either stimulated on inhibited in proportion to the concentration of the gradient substance.

That these parts of the chromosome physically unroll, in the sense of some process propagating different distances along the parts of the chromosomes where the hox genes are located, and (for some reason) this propagated change goes farther in cells located toward the rear end of the developing embryo.

A third category of explanation would be for transcription of particular hox genes in one cell to influence transcription of the same hox gene and other hox genes adjacent to it on the chromosomes. This would require that individual cells can somehow detect which hox genes are being transcribed and translated in adjacent cells, and compare those with which hox gene proteins it, itself, contains.

Concepts of pattern embryonic formation mostly belong to one or other of the following two categories:

I) Individual responses to **long-range signals**. ("Positional Information")
II) Cumulative results of many **close-range interactions**. (Turing, etc.)

Another example of a long-range signal method, suppose organ B needs to be between organ A and organ C; The long-range way to achieve this is for some one kind of signal to be sent to the cells at both locations, and for this one kind of signal to be a little stronger one place than another, for example stimulating cells to become A where the signal is strongest, for the cells to form organ B where the signal is intermediate in strength and to become organ C where this long range signal is weakest (or the reverse). Wolpert's theory of "Positional Information" is an example, but not the only example, of an embryological control mechanism based on long range signaling.

The two kinds of close range-signaling mechanisms that get attention **are reaction-diffusion systems** and **cellular automata**. (Although people who know about reaction-diffusion systems tend not to have heard of cellular automata, and vice versa).
Much more study should be made about the possibilities of complicated close range interaction.

Suppose that cells can detect which hox genes are being expressed in neighboring cells, and respond by activating or repressing their own hox genes according to rules that cause cells to express "the average" of their neighboring cells.

Imagine a vertical column of cells, expressing these Hox genes:

Cell A expresses	H_1
Cell B expresses	H_1
Cell C expresses	H_1 H_2
Cell D expresses	H_1 H_2
Cell E expresses	H_1 H_2 H_3
Cell F expresses	H_1 H_2 H_3
Cell G expresses	H_1 H_2 H_3 H_4
Cell H expresses	H_1 H_2 H_3 H_4
Cell I expresses	H_1 H_2 H_3 H_4 H_5
Cell J expresses	H_1 H_2 H_3 H_4 H_5
Cell K expresses	H_1 H_2 H_3 H_4 H_5 H_6
Cell L expresses	H_1 H_2 H_3 H_4 H_5 H_6
Cell M expresses	H_1 H_2 H_3 H_4 H_5 H_6 H_7

Cell N expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7
Cell O expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8
Cell P expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8
Cell Q expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9
Cell R expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9
Cell S expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9 H_{10}
Cell T expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9 H_{10}
Cell U expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9 H_{10} H_{11}
Cell V expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9 H_{10} H_{11}
Cell W expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9 H_{10} H_{11} H_{12}
Cell X expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9 H_{10} H_{11} H_{12}
Cell Y expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9 H_{10} H_{11} H_{12} H_{13}
Cell Z expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9 H_{10} H_{11} H_{12} H_{13}

The pattern above resembles the actual pattern of hox gene expression.

But an embryo could start out with the pattern below, with cell A locked into expressing hox gene 1

Cell A expresses H_1
Cell B expresses H_1
Cell C expresses H_1
Cell D expresses H_1
Cell E expresses H_1
Cell F expresses H_1
Cell G expresses H_1
Cell H expresses H_1
Cell I expresses H_1
Cell J expresses H_1
Cell K expresses H_1
Cell L expresses H_1
Cell M expresses H_1
Cell N expresses H_1
Cell O expresses H_1
Cell P expresses H_1
Cell Q expresses H_1
Cell R expresses H_1
Cell S expresses H_1
Cell T expresses H_1
Cell U expresses H_1
Cell V expresses H_1
Cell W expresses H_1
Cell X expresses H_1
Cell Y expresses H_1
Cell Z expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9 H_{10} H_{11} H_{12} H_{13}

And also with cell Z locked into expressing all 13 hox genes.

Cell A expresses H_1
Cell B expresses H_1
Cell C expresses H_1
Cell D expresses H_1
Cell E expresses H_1
Cell F expresses H_1
Cell G expresses H_1
Cell H expresses H_1
Cell I expresses H_1
Cell J expresses H_1
Cell K expresses H_1
Cell L expresses H_1
Cell M expresses H_1
Cell N expresses H_1
Cell O expresses H_1 This pattern will be created after a few rounds of “averaging”.
Cell P expresses H_1
Cell Q expresses H_1
Cell R expresses H_1
Cell S expresses H_1
Cell T expresses H_1
Cell U expresses H_1 H_2

Cell V expresses H_1 H_2 H_3
Cell W expresses H_1 H_2 H_3 H_4 H_5 H_6
Cell X expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8
Cell Y expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9 H_{10}
Cell Z expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9 H_{10} H_{11} H_{12} H_{13}

Eventually, after many cycles of "averaging", the pattern below would be produced:

Cell A expresses H_1
Cell B expresses H_1
Cell C expresses H_1 H_2
Cell D expresses H_1 H_2
Cell E expresses H_1 H_2 H_3
Cell F expresses H_1 H_2 H_3
Cell G expresses H_1 H_2 H_3 H_4
Cell H expresses H_1 H_2 H_3 H_4
Cell I expresses H_1 H_2 H_3 H_4 H_5
Cell J expresses H_1 H_2 H_3 H_4 H_5
Cell K expresses H_1 H_2 H_3 H_4 H_5 H_6
Cell L expresses H_1 H_2 H_3 H_4 H_5 H_6
Cell M expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7
Cell N expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7
Cell O expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8
Cell P expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8
Cell Q expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9
Cell R expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9
Cell S expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9 H_{10}
Cell T expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9 H_{10}
Cell U expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9 H_{10} H_{11}
Cell V expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9 H_{10} H_{11}
Cell W expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9 H_{10} H_{11} H_{12}
Cell X expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9 H_{10} H_{11} H_{12}
Cell Y expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9 H_{10} H_{11} H_{12} H_{13}
Cell Z expresses H_1 H_2 H_3 H_4 H_5 H_6 H_7 H_8 H_9 H_{10} H_{11} H_{12} H_{13}

Which is the pattern that actual developing embryos develop.

But the usual explanation for this pattern of hox gene expression is that long-range signals separately instruct each of the cells which hox genes to express

Imagine a class room full of very nearsighted people, who are extreme conformists. One example of their conformism is that they get their hair cut to match the average length of the hair of the people sitting next to them.

Then imagine that a woman in the back row has very long hair, but that the professor happens to be bald. So the conformists in the front row get "crew cuts", and those toward the back let their hair grow longer.

The eventual result will resemble a diffusion gradient, as if hair were being synthesized at the back, diffusing forward toward some kind of sink. But really it's the result of close range conformism.