Research

Open Access

The genomic response to 20-hydroxyecdysone at the onset of *Drosophila* metamorphosis Robert B Beckstead, Geanette Lam and Carl S Thummel

Address: Department of Human Genetics, Howard Hughes Medical Institute, University of Utah School of Medicine, Salt Lake City, UT 84112-5331, USA.

Correspondence: Carl S Thummel. E-mail: carl.thummel@genetics.utah.edu

Published: 21 November 2005

Genome Biology 2005, 6:R99 (doi:10.1186/gb-2005-6-12-r99)

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2005/6/12/R99

Received: 17 June 2005 Revised: 5 August 2005 Accepted: 20 October 2005

© 2005 Beckstead et al.; licensee BioMed Central Ltd.

This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: The steroid hormone 20-hydroxyecdysone (20E) triggers the major developmental transitions in *Drosophila*, including molting and metamorphosis, and provides a model system for defining the developmental and molecular mechanisms of steroid signaling. 20E acts via a heterodimer of two nuclear receptors, the ecdysone receptor (EcR) and Ultraspiracle, to directly regulate target gene transcription.

Results: Here we identify the genomic transcriptional response to 20E as well as those genes that are dependent on *EcR* for their proper regulation. We show that genes regulated by 20E, and dependent on *EcR*, account for many transcripts that are significantly up- or downregulated at puparium formation. We provide evidence that 20E and EcR participate in the regulation of genes involved in metabolism, stress, and immunity at the onset of metamorphosis. We also present an initial characterization of a 20E primary-response regulatory gene identified in this study, *brain tumor* (*brat*), showing that *brat* mutations lead to defects during metamorphosis and changes in the expression of key 20E-regulated genes.

Conclusion: This study provides a genome-wide basis for understanding how 20E and its receptor control metamorphosis, as well as a foundation for functional genomic analysis of key regulatory genes in the 20E signaling pathway during insect development.

Background

Small lipophilic hormones such as retinoic acid, thyroid hormone, and steroids control a wide range of biological pathways in higher organisms. These hormonal signals are transduced into changes in gene expression by members of the nuclear receptor superfamily that act as hormone-responsive transcription factors [1]. Although extensive studies have defined the molecular mechanisms by which nuclear receptors regulate transcription, much remains to be learned about how these changes in gene activity result in the appropriate biological responses during development.

Drosophila melanogaster provides a powerful model system for elucidating the molecular and genetic mechanisms of hormone action. Pulses of the steroid hormone 20-hydroxyecdysone (20E) act as critical temporal signals that direct each of the major developmental transitions in the *Drosophila* life cycle, including molting and metamorphosis [2].



Figure I

Validation of the temporal patterns of 20E-regulated gene expression as determined by microarray analysis. (a) Northern blot hybridizations adapted with permission from published data [27,28]. Arrow indicates *E74A* isoform. (b) Cluster analysis of microarray data derived from RNA samples isolated from staged wild-type animals. The colors for each time point represent the change in the level of expression relative to the average expression levels across all time points for that gene, with dark blue indicating the lowest level of expression and red indicating the highest level, as depicted on the bottom. The numbers at the top indicate hours relative to pupariation, with green bars representing the peaks of 20E titer.

A high titer pulse of 20E at the end of the third larval instar triggers puparium formation, initiating metamorphosis and the prepupal stage of development. A second 20E pulse approximately 10 hours after pupariation triggers adult head eversion and marks the prepupal-to-pupal transition. Our current understanding of the molecular mechanisms of 20E action in insects derives from detailed characterization of the puffing patterns of the giant larval salivary gland polytene chromosomes [3-6]. These studies exploited an organ culture system that allows the use of defined hormone concentrations as well as the addition of cycloheximide to distinguish primary responses to the 20E signal [5,6]. The puffing studies revealed that 20E acts, at least in part, through a two-step regulatory cascade. The hormone directly induces approximately six early puff genes [7]. The protein products of these genes were proposed to repress their own expression as well as induce many secondary-response late puff genes that, in turn, were assumed to direct the appropriate biological responses to the hormone.

The identification and characterization of a 20E receptor, along with several early and late puff genes, has supported and extended this hierarchical model of 20E action. Like vertebrate hormones, 20E regulates gene expression by binding to a nuclear receptor heterodimer, consisting of the ecdysone receptor (EcR) and Ultraspiracle (USP), which are orthologs of the vertebrate LXR and RXR receptors, respectively [8]. Several early puff genes have been identified, including the



Figure 2 (see legend on next page)

Figure 2 (see previous page)

Microarray results for *EcRi* and 20E organ cultures experiments. (a) Graphic depiction of the number of genes upregulated (blue) or downregulated (yellow) in *EcRi* late third instar larvae or prepupae at the times indicated. (b) Graphic depiction of the number of 20E-regulated genes or 20E primary-response genes that are either upregulated (blue) or downregulated (yellow) in third instar larval organ culture. (c) Venn diagram depicting the overlap between all *EcR*-regulated genes and the combined 20E-regulated genes and 20E primary-response genes. (d) Cluster diagram depicting the temporal expression pattern of the 479 genes in the 20E-final set, divided into those genes that are upregulated by 20E (above) or downregulated by 20E (below). Times are shown in hours relative to puparium formation, and colors are as described in Figure 1b.

Broad-Complex (BR-C) and E74 [9,10]. As predicted by the puffing studies, these genes encode transcription factors that directly regulate late puff gene expression [11,12] and are essential for appropriate biological responses to 20E [13,14]. Other studies, however, have shown that not all early puffs encode transcriptional regulators. These include a calcium binding protein encoded by E63-1 [15] and the E23 ABC transporter gene [16]. In addition, a molecular screen identified fifteen new 20E primary-response genes, only two of which correspond to early puff loci, suggesting that the hormone triggers a much broader transcriptional response than is evident from the puffing pattern of the salivary gland polytene chromosomes [17]. Similarly, the isolation of late puff genes has demonstrated that some of these presumed effectors may function in a regulatory capacity, such as the CDKlike protein encoded by L63 [18].

Several papers have used microarrays to identify genes that change their expression at the onset of metamorphosis [19-21]. Although critical for understanding the dramatic switches in gene expression that occur at this stage, these studies are restricted to developmental analysis of staged tissues or animals, with no direct links to 20E signaling. Increasing evidence indicates that other hormones and receptors may contribute to the complex developmental pathways associated with metamorphosis [8,22,23]. In addition, some transcripts are induced at puparium formation independently of 20E or its receptor [24]. It thus remains unclear to what extent 20E and EcR contribute to the global reprogramming of gene activity that occurs at the early stages of metamorphosis.

In this study, we use larval organ culture in combination with microarray technology to identify genes regulated by 20E alone or 20E in the presence of cycloheximide [5,6]. We also examine the effects of disrupting *EcR* function on the global patterns of gene expression at the onset of metamorphosis, and use these data to refine our lists of 20E-regulated genes. The top 20E-regulated genes described here include many of the key genes identified by puffing studies, validating our approach. We also identify many new genes that are part of the 20E/EcR regulatory cascade and define roles for EcR in the regulation of stress, immunity, and metabolism at the onset of metamorphosis. Finally, we characterize one 20E primary-response target in more detail - the brat gene, which encodes a translational regulator [25,26]. We show that brat mutants display defects at the onset of metamorphosis and mis-regulate key 20E target genes, consistent with a disruption in 20E signaling. This work provides a genomic foundation for defining the roles of 20E and EcR in controlling insect development.

Results and discussion Most genes that change expression at pupariation require EcR function

To identify genes that alter their expression in synchrony with the late third instar and prepupal pulses of 20E, RNA was isolated from *w*¹¹¹⁸ animals staged at -18, -4, 0, 2, 4, 6, 8, 10, and 12 hours relative to pupariation, labeled, and hybridized to Affymetrix Drosophila Genome Arrays. The sensitivity and accuracy of the array data were determined by comparing the expression patterns of known 20E-regulated genes with previously published developmental Northern blot data [27,28]. A subset of this analysis, depicted in Figure 1, reveals that the temporal expression pattern of key regulatory genes - EcR, usp, E74A, DHR3, FTZ-F1, and DHR39 - are faithfully reproduced in the temporal arrays, as well as the 20E-regulated switch from Sgs glue genes to L71 late genes in the larval salivary glands, and the expression of representative IMP and Edg genes in the imaginal discs and epidermis. This comparison demonstrates that the microarrays accurately reflect the temporal patterns of 20E-regulated gene expression at the onset of metamorphosis and have sufficient sensitivity to detect rare transcripts such as EcR and E74A.

EcR mutants die during early stages of development, complicating their use for studying receptor function during metamorphosis. To circumvent this problem, we employed a transgenic system that allows heat-induced expression of double-stranded RNA corresponding to the EcR common region to disrupt *EcR* function at puparium formation (*EcRi*) [29]. RNA was harvested for array analysis from *EcRi* animals staged at -4, 0, and 4 hours relative to pupariation. All EcRi animals formed arrested elongated prepupae, consistent with an effective block in 20E signaling and highly reduced EcR protein levels (Figures 2 and 3c in [29]). Data obtained from these arrays were compared to our array data from control animals at the same stages of development to identify EcRdependent genes. The initial effect of EcR RNA interference RNA (RNAi) is significant upregulation of gene expression in late third instar larvae, followed by a switch at puparium formation such that the majority of genes are not properly induced (Figure 2a). These data are consistent with genetic studies of usp that define a critical role for this receptor in repressing ecdysone-regulated genes during larval stages



Figure 3

Temporal expression patterns of *EcR*-dependent genes that are regulated by (a) starvation, (b) stress, or (c) infection. Upregulated (up) and downregulated (down) genes are labeled, hours are relative to puparium formation, which is marked by a black line, and colors are as described in Figure 1b.

[30], and provide further evidence that one essential function for the EcR-USP heterodimer is to prevent premature maturation through the repression of select 20E target genes during larval stages.

A total of 4,188 genes change their expression at least 1.5-fold in at least one time point in EcRi animals (Additional data file 1), suggesting that almost a third of all genes require *EcR*, either directly or indirectly, for their proper regulation at the onset of metamorphosis. This number is consistent with the 2,268 genes that have been reported to change their expression at pupariation in one of five tissues examined: midgut, salivary gland, wing disc, epidermis, and central nervous system [20]. It is also similar to the 4,042 genes that change their expression at least 1.5-fold at pupariation in our temporal arrays. Of these 4,042 genes, 2,680 are affected in EcRi animals, supporting the proposal that EcR plays a major role in coordinating transcriptional responses at the onset of metamorphosis. Not all genes that change their expression at pupariation, however, are dependent on EcR. Several such transcripts were selected for validation by Northern blot hybridization (Additional data file 3). This is consistent with an earlier microarray study of EcR-regulated genes in the larval midgut [20]. This study found that of 955 genes that change their expression in wild-type midguts at the onset of metamorphosis, 672 genes are affected by an *EcR* mutation while 283 genes are unaffected, close to the proportion of EcR-independent genes identified by our study. This is also consistent with earlier studies that indicate that other signaling pathways are active at this stage in development. For example, the miR-125 and let-7 microRNAs are dramatically induced at puparium formation, in tight temporal synchrony with the 20E primary-response E74A mRNA, but do so in a manner that is independent of either 20E or EcR [24]. Similarly, α -ecdysone, the immediate upstream precursor of 20E,

has critical biological functions [23,31,32], can activate the DHR38 nuclear receptor [22], and can induce genes in *Drosophila* third instar larvae that are distinct from those that respond to 20E (RBB, GL and CST, unpublished results). The sesquiterpenoid juvenile hormone can also function with 20E to direct specific transcriptional responses during early metamorphosis [33-35]. The results of the study described here, however, indicate that most genes that change their expression at the onset of metamorphosis do so in an *EcR*-dependent fashion, and pave the way for future studies that integrate these responses with those of other signaling pathways.

Identification of 20E-regulated genes in cultured larval organs

To identify 20E-regulated genes, wandering third instar larvae were dissected and their organs cultured in the presence of either no hormone, 20E alone, cycloheximide alone, or 20E plus cycloheximide for 6 hours. RNA extracted from these samples was analyzed on Affymetrix Drosophila Genome Arrays. Comparison of the no hormone and 20E-treated datasets led to the identification of 20E-regulated genes, while comparison of the cycloheximide dataset with data derived from organs treated with 20E and cycloheximide led to the identification of a set of genes we refer to as 20E primaryresponse genes. In comparing these datasets, it is important to note that cycloheximide treatment alone can stabilize preexisting mRNAs and thus mask their induction by 20E [6,17,36]. These transcripts would not be identified by our experiments. In addition, some 20E-inducible genes are expressed at higher levels in the absence of protein synthesis, due to the lack of 20E-induced repressors [6,17]. The addition of cycloheximide thus provides a means of detecting 20E-regulated transcripts that might otherwise be missed. In this study, 743 20E-regulated genes were identified (Figure 2b), with 555 genes responding to 20E alone, 345 genes Table I

Microarrays		<i>EcRi</i> ↑ (n = 634)	<i>EcRi</i> ↓ (n = 924)	20E↑ (n = 411)	20E↓ (n = 68)
A	midgut; $EcR\uparrow$ (n = 371)	19 (5.8E-01)	120 (8.4E-91)*	65 (1.3E-63)*	4 (9.7E-02)
В	midgut; $EcR\downarrow$ (n = 292)	113 (1.2E-176)*	12 (8.2E-02)	3 (5.0E-02)	10 (3.1E-13)*
С	Steroid CD↑ (n = 916)	34 (2.1-01)	192 (5.3E-73)*	125 (1.7E-87)*	2 (2.3E-01)
D	Steroid CD \downarrow (n = 931)	53 (8.0E-02)	157 (9.3E-39)*	59 (2.3E-10)*	16 (2.3E-08)
E	Starv Sug↑ (n = 718)	120 (3.1E-58)*	87 (I.IE-09)	47 (4.5E-09)	12 (2.9E-06)
F	Starv Sug↓ (n = 667)	55 (4.3E-06)	90 (7.7E-13)*	32 (4.9E-03)	10 (1.5E-04)
G	Stress↑ (n = 253)	33 (5.5E-11)*	34 (I.IE-05)	12 (8.7E-02)	9 (1.4E-12)*
н	Stress↓ (n = 310)	87 (4.7E-90)*	25 (3.0E-01)	9 (9.7E-01)	3 (2.2E-01)
I	Immune↑ (n = 221)	56 (I.IE-50)*	29 (8.8E-05)	14 (2.6E-03)	13 (3.5E-31)*
J	Immune↓ (n = 148)	40 (7.5E-40)*	17 (1.7E-02)	4 (8.6E-01)	3 (6.8E-03)

This table depicts a comparison of *EcRi* (columns I and 2) and 20E-final (columns 3 and 4) microarray data with five other gene sets: rows A and B, *EcR*-dependent genes expressed in the larval midgut [20]; rows C and D, cell death genes expressed in the larval salivary gland [37]; rows E and F, genes that change expression in response to starvation or a sugar diet [45]; rows G and H, genes that change expression in response to either paraquot, tunicamycin, or H_2O_2 [43]; and rows I and J, genes that change expression in response to bacterial insult [44]. Each gene set is divided into upregulated or downregulated genes as represented by the arrows, with the number of genes in each dataset represented by '(n =)'. The first number in each cell represents the number of overlapping genes between the two datasets being compared. The numbers within the parentheses in each cell represent a *p* value based on the χ^2 test that accounts for the differences between the observed and expected numbers. Correlations discussed in the text are marked with an asterisk (all those showing a *p* value $\leq E-10$).

responding to 20E in the presence of cycloheximide (Additional data file 2), and 159 genes overlapping between these two datasets.

Comparison of the 20E-regulated genes to those genes that require EcR for their proper regulation at the onset of metamorphosis led to a final list of 20E-regulated, EcR-dependent genes. Only those genes that are upregulated by 20E in culture and downregulated in at least one of the EcRi time points, or downregulated by 20E in culture and upregulated in at least one of the *EcRi* time points, were considered for further analysis, leading to the identification of 479 genes (20E-final; Figure 2c). As depicted in Figure 2d, the majority of 20E-final genes that are upregulated by 20E are induced in -4 hour late larvae and/or early prepupae, in apparent response to the late larval 20E pulse, while many genes downregulated by 20E are repressed at these times. The downregulated 20E-final genes that peak in 4 to 6 hour prepupae could be repressed by 20E and thus expressed during this interval of low 20E titer.

We compared our *EcR*-dependent genes and the 20E-final gene set to data from two microarray studies that examined 20E-regulated biological responses - either *EcR*-dependent genes expressed in the larval midgut at pupariation [20] (Table 1, rows A and B), or changes in gene expression that occur during 20E-induced larval salivary gland cell death [37] (Table 1, rows C and D). As expected, many genes that are normally downregulated in the midgut at pupariation are upregulated in our *EcRi* gene set (113 genes; Table 1, row B), and genes that are normally upregulated in the midgut at the midgut at

pupariation are downregulated in our *EcRi* gene set (120 genes; Table 1, row A). Similarly, we see significant overlaps between our 20E-final set and midgut genes that change their expression at pupariation (65 genes upregulated and 10 genes downregulated; Table 1, rows A and B). Statistically significant overlaps were also observed with genes that change their expression during salivary gland cell death, consistent with a critical role for 20E in directing this response [37] (Table 1, rows C and D). These correlations validate our datasets and support the conclusion that our results represent 20E responses in multiple tissues at the onset of metamorphosis.

Those genes that are upregulated by 20E twofold or higher and dependent on EcR are listed in Table 2. An examination of this list reveals several known key mediators of 20E signaling during development. These include three classic ecdysone-inducible puff genes, E74A, E75, and E78 [7], as well as Kr-h1, which encodes a family of zinc finger proteins required for metamorphosis [38], the DHR3 nuclear receptor gene [39], and Cyp18a1 [17]. Expanding our list by including all 20E-regulated genes (Additional data file 2), results in the identification of the DHR39, DHR78, and FTZ-F1 nuclear receptor genes, as well as the L71 (Eip71E) late genes, IMP-E2, IMP-L3, Fbp-2, Sgs-1, urate oxidase, and numerous genes identified in other studies as changing their expression at the onset of metamorphosis [20,21,40]. The identification of well-characterized 20E-regulated genes within our datasets suggests that the other genes in these lists are also likely to function in 20E signaling pathways, and thus provide a foundation to extend our understanding of 20E action in new directions.

Table 2

20E-induced, EcR-dep	R-dependent genes					
Gene	20E	20E primary	<i>EcRi -</i> 4 h	<i>EcRi</i> 0 h	EcRi 4 h	Function
Eip74A*	11.0	12.6	-4.9	-4.5	-2.0	Transcription factor
Drosomycin-2	8.I		-4.5	-7.3	-14.5	Defense response
CG14664	4.7		-16.3	-22.9		n/o
GST-E3	3.7	3.2		-1.9		Glutathione transferase
Eip78C*	3.6	3.3	-6.2	-6.9	-1.9	Nuclear receptor
CG10444	3.4		-4.1	-7.1	-2.1	Na ⁺ -dependent multivitamin transporter
Punch	3.0	2.4	-3.1	-8.1	-2.2	GTP cyclohydrolase I
CG5171	3.0	2.7	-1.7	-1.5	-3.5	Trehalose-phosphatase
CG2444	3.0		-2.8	-2.4	2.3	n/o
Cyp18a1*	2.9	8.3	-2.3		-1.7	Cytochrome P450
CG5391	2.7		-5.2		3.1	n/o
MESR3	2.6		-1.5	-3.0	-1.6	Ras signaling
CG3714	2.6	2.3		-1.7		Nicotinate phosphoribosyltransferase
rdgBbeta	2.5	4.6		-2.2	-2.4	Phosphatidylinositol transporter
CG8501	2.4			-1.9	-1.8	n/o
l(3)00305	2.4	1.9		-1.6		Serine/threonine kinase
CG11737	2.4		-2.3	-2.0		n/o
Cabut	2.3	5.9	-1.9			Transcription factor
CG17834	2.3	1.6		-2.1	-4.4	n/o
Sox box protein 14	2.3	2.9	-1.5	-1.7		Transcription factor
JIL-I	2.2				-3.0	Protein serine/threonine kinase
CG18212	2.2			-2.1	-1.8	n/o
CG13252	2.2	2.0	-2.1	-2.4	-2.7	n/o
raspberry	2.2	1.7	-1.9	-2.2	-1.8	IMP dehydrogenase
supernumerary limbs	2.2			-3.1	-2.6	Ubiquitin-protein ligase
CG16995	2.1	1.5			-2.8	Defense response
lamina ancestor	2.1	2.9	-1.6	-3.1	-2.4	n/o
CG33090	2.1			-3.9	-1.5	n/o
CG7701	2.1			-1.6		n/o
rhea	2.1			-1.9	-2.2	Actin binding
Ptp52F	2.1	2.2	-2.4	-1.6		Protein tyrosine phosphatase
CG4822	2.1	1.4		-3.1	-2.2	ATPase activity
big brain	2.1				-2.4	Water channel
brain tumor	2.0	2.9		-1.7		Translation regulator
CG11529	2.0		-7.1	-2.2	2.0	Serine-type endopeptidase
CG11509	2.0	1.9			-1.6	n/o
CG3348	2.0	4.5	-3.0	-3.1	-1.7	n/o
CG9989	2.0		-1.5	-2.2	-2.5	n/o
blue cheese	1.9	2.0		-2.0	-1.6	Intracellular protein transport
CG14073	1.9	2.4			-1.5	Ankyrin-like
kruppel homolog I*	1.9	3.5		-1.8	-2.2	Transcription factor
CRMP	1.8	2.6		-2.3	-3.2	Dihydropyrimidinase
CG9801	1.8	3.1		-4.5	-5.7	n/o
CG8483	1.8	5.4	-3.0	-2.8		Defense response
CG9005	1.8	2.4	-1.8	-1.8	-1.8	Cell adhesion
vrille	1.6	3.5			-1.6	Transcription factor
CG1342	1.6	2.4		-2.6	-8.5	Proteinase inhibitor I4, serpin
DHR3*	1.6	2.3			-2.9	Nuclear receptor

20E-induced, EcR-dependent genes						
Eip75B*	1.5	2.8	-1.6			Nuclear receptor
hairy		4.4		-1.8	-1.6	Transcription factor
CG9192		3.3			-1.7	n/o
CG5249		3.0			-1.5	Transcription factor
black		2.7	-3.0	-7.0		Glutamate decarboxylase 2
CG2016		2.1	-2.0	-2.0	-2.0	n/o
CG12539		2.1	-1.7		-1.6	Glucose dehydrogenase
CG8788		2.0	-1.5	-2.5	-1.8	n/o

Table 2 (Continued)

Genes that show at least a twofold induction with either 20E alone (20E), or 20E + cycloheximide (20E primary) are listed in the order of their foldinduction by 20E alone. Downregulation of these genes upon *EcR* RNAi is shown for each time point, -4, 0, or 4 hours relative to puparium formation. Function is inferred from gene ontology on FlyBase. [40]. Asterisks denote previously identified 20E-regulated genes. n/o, no ontology.

Immunity, stress-response, and starvation genes are regulated by 20E at pupariation

In an effort to identify biological pathways that might respond to 20E at the onset of metamorphosis, we compared our EcRi and 20E-final datasets with published microarray studies of circadian rhythm, starvation, stress, and immunity [41-45]. No statistically significant overlaps were seen with the circadian rhythm gene sets examined; however, we did observe significant overlaps with genes that are expressed during starvation, stress, or an innate immune response. For the starvation response, we examined genes that change their expression upon starvation for 4 hours or starvation in the presence of sugar for 4 hours [45]. We observed 120 genes induced under these conditions that are upregulated in EcRi animals, and 90 genes that are repressed upon starvation and downregulated in EcRi animals (Table 1, rows E and F). As shown in Figure 3a, the starvation-regulated genes are part of an EcR-dependent switch that occurs at puparium formation, where many of the induced genes are normally downregulated at puparium formation, and many starvation-repressed genes are upregulated at puparium formation. These genes include eight members of the cytochrome P450 family, three triacylglycerol lipase genes, α-trehalose-phosphate synthase, and a fatty-acid synthase gene that are downregulated at the onset of metamorphosis, while lipid storage droplet-1, pumpless, a UDP-galactose transporter, a lipid transporter, and phosphofructokinase are upregulated at this stage. Similarly, genes that change their expression in response to oxidative or endoplasmic reticulum stress [43] are significantly upregulated in EcRi animals at puparium formation (Table 1, rows G and H), reflecting their normal coordinate downregulation at puparium formation (Figure 3b), and demonstrating that this response is mediated by EcR. Within the 87 genes that overlap between the downregulated stress response genes and the upregulated EcR-dependent genes, we identified 14 of the 17 Jonah genes that encode a family of coordinately regulated midgut-specific putative proteases [46]. Six genes that encode trypsin family members are also within this gene set, indicating that many peptidase family members are regulated

by EcR. Taken together with the data on *EcR*-regulated starvation genes, these results indicate that EcR plays a central role in controlling metabolic responses at pupariation, directing the change from a feeding growing larva to an immobile non-feeding pupa.

Genes that change their expression upon microbial infection [44,47] are also significantly upregulated in *EcRi* animals at puparium formation (Table 1, rows I and J), and coordinately downregulated at pupariation (Figure 3c). Interestingly, we identified both the Toll ligand-encoding gene dorsal and the key Toll effector gene spätzle as downregulated at the onset of metamorphosis in a *EcR*-dependent manner, suggesting that central regulators of the Toll-mediated immune response pathway are under EcR control [48]. In addition, well studied immune response genes are downregulated by 20E, including Cecropin C, Attacin A, Drosocin, Drosomycin, and Defensin (Additional data file 2). These observations indicate that many metabolic and immunity-regulated genes are part of the genetic program directed by 20E at the onset of metamorphosis, and that these genes are normally coordinately downregulated at puparium formation in an EcR-dependent manner.

Identification of novel 20E primary-response regulatory genes

We selected all potential transcriptional and translational regulators from the list of most highly induced 20E primaryresponse genes that are *EcR*-dependent (Table 2) and not yet implicated in 20E signaling pathways, identifying seven genes: *sox box protein 14* (*sox14*), *cabut*, *CG11275*, *CG5249*, *vrille*, *hairy*, and *brain tumor* (*brat*). Northern blot hybridization was used to validate the transcriptional responses of these genes to 20E (Figure 4a). All seven genes are induced by 20E in larval organ culture, with *CG5249* displaying a very low level of expression and *hairy* showing only a modest approximately twofold induction. Several transcripts are increased upon treatment with cycloheximide alone, consistent with its known role in stabilizing some mRNAs [6,17,36].



Figure 4

Validation of seven 20E primary-response regulatory genes. (a) Northern blot analysis of RNA samples isolated from organ cultures treated with either 20E alone, 20E plus cycloheximide (20E+Cyc), or cycloheximide (Cyc) alone, for 0, 2, or 6 hours. (b) Temporal expression patterns of depicted genes with hours shown relative to puparium formation. Green bars represent the peak 20E titers. Colors in the cluster analysis are as described in Figure 1b. Hybridization to detect *E74A* and *rp49* mRNAs was included as a control.

Addition of 20E and cycloheximide, however, resulted in higher levels of transcript accumulation, similar to the response seen when E74A is used as a control [6,49]. Their temporal patterns of expression at the onset of metamorphosis also reveal brief bursts of transcription that correlate with the 20E pulses that trigger puparium formation and adult head eversion (Figure 4b). These seven genes thus appear to represent a new set of 20E primary-response regulatory genes that could act to transduce the hormonal signal during metamorphosis.

brat is required for genetic and biological responses to 20E during metamorphosis

We examined roles for *brat* during metamorphosis because, unlike the other six 20E primary-response genes described above, a *brat* mutant allele is available (*brat*^{ko6028}) that allows an assessment of its functions during later stages of development [25,50]. The *brat*^{ko6028} *P*-element maps to the fourth exon of the *brat* gene. Precise excisions of this transposon result in viable, fertile animals, demonstrating that the transposon is responsible for the mutant phenotype [25]. Lethal phase analysis of *brat*^{ko6028} mutants revealed that 61% of the animals survive to pupariation, with the majority of these animals pupariating 1 to 2 days later than their heterozygous siblings (n = 400). Of those mutants that pupariated, 11% died as prepupae, 8% died as early pupae, 46% died as pharate adults, and the remainder died within a week of adult eclosion. Phenotypic characterization of bratko6028 mutant prepupae and pupae revealed defects in several ecdysone regulated developmental processes, including defects in anterior spiracle eversion (29%; Figure 5b-d), malformed pupal cases (15%, Figure 5b-d), and incomplete leg and wing elongation (12%). Northern blot hybridization of RNA isolated from staged bratko6028 mutant third instar larvae (Figure 5e, -18 and -4 hour time points) or prepupae (Figure 5e, 0 to 12 hours) revealed a disruption in the 20E-regulated transcriptional hierarchy. In wild type animals, brat mRNA is induced in late third instar larvae and 10 hour prepupae, similar to the temporal profile determined by microarray analysis (Figures 4b and 5e), with reduced levels of brat mRNA in bratko6028 mutants, consistent with it being a hypomorphic allele [25]. βFTZ-F1 is unaffected by the brat mutation in mid-prepupae, while E74 mRNA is reduced at 10 hours after pupariation (Figure 5e). BR-C, E93, EcR, DHR3, and L71-1 are expressed at higher levels in late third instar larvae and early prepupae (Figure 5e), with significant upregulation of BR-C. In addition, the smallest BR-C mRNA, encoding the Z1 isoform, is under-expressed in brat mutant prepupae (Figure 5e). It is

unlikely that *brat* exerts direct effects on transcription since it encodes a translational regulator [26]. Nonetheless, these effects on 20E-regulated gene expression are consistent with the late lethality of *brat^{ko6o28}* mutants. In particular, the *rbp* function provided by the BR-C Z1 isoform is critical for developmental responses to 20E, and overexpression of BR-C isoforms can lead to lethality during metamorphosis [13,51]. Thus, not only are the *brat* mutant phenotypes consistent with it playing an essential role during metamorphosis, but it may exert this function through the regulation of key 20Einducible genes. Efforts are currently underway to address the roles of the remaining six new 20E primary-response regulatory genes in transducing the hormonal signal at the onset of metamorphosis.

Conclusions

The classic studies of the giant larval salivary gland polytene chromosomes established a new paradigm for the mechanisms of steroid hormone action, raising the exciting possibility that these hormones could act directly on the nucleus, triggering a complex regulatory cascade of gene expression [7,52]. Although subsequent molecular experiments confirmed and significantly expanded this hierarchical model of 20E action, no studies to date have addressed the genomic effects of 20E on gene regulation or the global effects of EcR on gene expression at the onset of metamorphosis. The work described here provides a new basis for our understanding of 20E signaling, returning to the genome-wide level of the original puffing studies, but identifying individual genes that act in this pathway. Much as earlier studies of puff genes provided a foundation for our understanding of steroid hormone action, we envision that future molecular and genetic characterization of 20E-regulated, EcR-dependent genes will expand our understanding of 20E action and insect maturation in new directions.

Materials and methods

Animals, staging, and phenotypic analysis

 w^{1118} animals were used for phenotypic, array, and Northern blot studies. $brat^{ko6o28}/CyO$, kr-*GFP* was used to analyze bratfunction. Third instar larvae were staged by the addition of 0.05% bromophenol blue to the food as previously described [53], or synchronizing animals at pupariation. $brat^{ko6o28}$ animals were identified by the loss of the kr-*GFP* marker associated with the *CyO* balancer chromosome. For *EcR* RNAi, *hs*-*EcRi-11* third instar larvae were heat-treated twice at 37°C, each time for 1 hour, at 24 hours and 18 hours prior to pupariation, as described [29]. RNA was harvested for microarray analysis at -4, 0, and 4 hours relative to pupariation from three independent collections of animals for each time point.

Organ culture

Partial blue gut third instar larvae were staged by the addition of 0.05% bromophenol blue to the food, as previously described [53]. Eight animals were dissected in each well of a nine-well glass dish (Corning, Corning, NY, USA) and cultured in approximately 100 µl oxygenated Schneiders Drosophila Medium (Invitrogen, Carlsbad, CA, USA) at 25°C. Cultures were incubated in a styrofoam box under a constant flow of oxygen. Following an initial incubation of 1 hour, the medium was removed and replaced with either fresh Schneiders Drosophila Medium (no hormone), medium plus $8.5 \times$ 10⁻⁵ M cycloheximide (Sigma-Aldrich, St. Louis, MO, USA), medium plus 5×10^{-6} M 20-hydroxyecdysone (Sigma), or medium plus cycloheximide and 20E, each for 6 hours at 25°C. Organs were collected and RNA was extracted as described below. All experiments were done in triplicate and harvested separately for microarray analysis.

Microarray and cluster analysis

All experiments for microarray analysis were performed independently, in triplicate, to facilitate statistical analysis. Total RNA was isolated using TriPure (Roche, Indianapolis, IN, USA) followed by further purification with RNAeasy columns (Qiagen, Valencia, CA, USA). Probe labeling, hybridization to Affymetrix GeneChip® Drosophila Genome Arrays (Affymetrix, Santa Clara, CA, USA), and scanning, were performed by the University of Maryland Biotechnology Institute Microarray Core Facility. dChip1.2 was used to normalize the raw data and determine gene expression values [54]. Statistically significant changes between sample sets were identified using significance analysis of microarray (SAM) with a delta value to give a <10% false discovery rate [55]. Further analysis and comparisons between datasets were performed using Access (Microsoft Corporation, Redmond, WA, USA). Cluster analysis was performed using dChip1.2. A cutoff of 1.3-fold change in expression level was used to restrict the 20E-regulated gene set (organ culture data), with a 1.5-fold cutoff for the EcRi gene sets. These fold cutoffs were chosen in order to restrict the datasets to those genes that are most significantly affected by 20E and *EcRi*. The lower fold cutoff for the organ culture data reflects the observation that 20E responses tend to be reduced in organ culture when compared to the intact animal [17] (unpublished results). This is, most likely, due to

Figure 5 (see following page)

Mutations in *brat* lead to defects in genetic and biological responses to 20E. (a) Control w^{1118} pharate adult. (b-d) Representative *brat*^{k06028} mutant animals. (e) Northern blot analysis of w^{1118} and *brat*^{k06028} mutants staged in hours relative to pupariation. Blots were probed to detect *brat* mRNA and transcripts from seven different 20E-regulated genes. Hybridization to detect *rp49* mRNA was included as a control for loading and transfer.



the stress of dissection and *in vitro* culture. Microarray data from this study can be accessed at the National Center for Biotechnology Information Gene Expression Omnibus website [56], with accession numbers as follows: GSE3057 for the temporal array series; GSE3060 for the organ culture array series; and GSE3069 for the *EcRi* array series.

Comparisons were made between our microarray datasets and previously published microarray datasets using Microsoft Access. Each dataset was split into genes that are either upregulated or downregulated, represented by up or down arrows in Table 1. The EcRi datasets were constrained to at least a twofold change in expression level in order to reduce the 4,188 genes to more manageable numbers, resulting in 634 upregulated genes and 924 downregulated genes (Table 1). For the stress-regulated genes, only those genes that show at least a twofold change in expression with either paraquot, tunicamycin, or H₂O₂ treatment were used for comparison [43]. The 400 immune genes were split into 221 genes that showed consistent upregulation and 148 genes that showed consistent downregulation, eliminating 31 genes that showed an inconsistent profile under the conditions tested [44,47]. All other microarray datasets used for our comparison studies were as published (see Table 1 legend).

Northern blot hybridizations

Total RNA was isolated using Trizol (Gibco) from staged animals or organ cultures, fractionated by formaldehyde gel electrophoresis, transferred to nylon membranes, and probed with radioactively labeled probes [27]. To facilitate comparisons, w^{1118} blots and $brat^{k06028}$ blots where probed, washed, and exposed together. rp49 was used as a loading control. Probes were generated by PCR from genomic DNA using the following pairs of primers: hairy forward (F), 5'CAAATTGGAAAAGGCCGACA3'; hairy reverse (R), 5'AGAGAAACCCTAAGCGGCTT3'; cabut F, 5'CTCTTCTAGTAGCCAAGACG3'; cabut R. 5'GAGATTGGTTCTGATGCTGC3'; sox box protein 14 F, 5'TCAGCAAGGACGATAAGCAGC3'; sox box protein 14 R, F, 5'AGCTCCGTTGTTATCGTGTGC3'; CG5249 R, 5'ACGATGTGGATCCTGAGACG3'; CG5249 F, 5'GCTCCATCATCAGCATGTGC3'; CG11275 5'ACACAGATTGCGTGTTCCACG3'; R, CG11275 F, 5'TGGACAACGTGACTCCATACG5'; brain tumor 5'TCTCCACGAACTGGAGAACG3': brain R, tumor F, 5'TGATGGTGTGACTGTTGGTGG3'; vrille 5'ACAGTTGTTGGCATCGCTGC3'; vrille R, 5'GACAACAAGAAGGACGAGAGC3'.

Additional data files

The following additional data are available with the online version of this paper. Additional data file 1 lists the 4,188 genes that change their expression \geq 1.5-fold or \leq -1.5-fold in at least one time point in *EcRi* animals. Additional data file 2 lists 20E-regulated and 20E primary-response genes. Addi-

tional data file 3 is a figure showing Northern blot analysis of RNA samples isolated from w^{1118} or *EcRi* animals staged at - 4, 0, or 4 hours relative to pupariation.

Acknowledgements

We thank A Godinez for assistance with microarray hybridizations and data collection, K King-Jones for help with microarray analysis, data comparisons and valuable discussions, and K Baker and M Horner for critical comments on the manuscript. RBB is a Research Associate, GL is a Research Specialist, and CST is an Investigator with the Howard Hughes Medical Institute.

References

- Chawla A, Repa JJ, Evans RM, Mangelsdorf DJ: Nuclear receptors and lipid physiology: opening the X-files. Science 2001, 294:1866-1870.
- Riddiford LM: Hormones and Drosophila development. In The Development of Drosophila melanogaster Edited by: Bate M, Arias AM. New York: Cold Spring Harbor Press; 1993:899-939.
- 3. Clever U: Actinomycin and puromycin: Effects on sequential gene activation by ecdysone. *Science* 1964, 146:794-795.
- Becker HJ: [The puffs of salivary gland chromosomes of Drosophila melanogaster. Part I. Observations on the behaviour of a typical puff in the normal strain and in two mutants, giant and lethal giant larvae.] Chromosoma 1959, 10:654-678.
- Ashburner M: Patterns of puffing activity in the salivary gland chromosomes of Drosophila. VI. Induction by ecdysone in salivary glands of D. melanogaster cultured in vitro. Chromosoma 1972, 38:255-281.
- Ashburner M: Sequential gene activation by ecdysone in polytene chromosomes of Drosophila melanogaster. II. The effects of inhibitors of protein synthesis. Dev Biol 1974, 39:141-157.
- Ashburner M, Chihara C, Meltzer P, Richards G: Temporal control of puffing activity in polytene chromosomes. Cold Spring Harbor Symp Quant Biol 1974, 38:655-662.
- Riddiford LM, Cherbas P, Truman JW: Ecdysone receptors and their biological actions. Vitam Horm 2000, 60:1-73.
- DiBello PR, Withers DA, Bayer CA, Fristrom JW, Guild GM: The Drosophila Broad-Complex encodes a family of related proteins containing zinc fingers. Genetics 1991, 129:385-397.
- Burtis KC, Thummel CS, Jones CW, Karim FD, Hogness DS: The Drosophila 74EF early puff contains E74, a complex ecdysone-inducible gene that encodes two ets-related proteins. Cell 1990, 61:85-99.
- Urness LD, Thummel CS: Molecular analysis of a steroidinduced regulatory hierarchy: The Drosophila E74A protein directly regulates L71-6 transcription. EMBO J 1995, 14:6239-6246.
- Crossgrove K, Bayer CA, Fristrom JW, Guild GM: The Drosophila Broad-Complex early gene directly regulates late gene transcription during the ecdysone-induced puffing cascade. Dev Biol 1996, 180:745-758.
- Kiss I, Beaton AH, Tardiff J, Fristrom D, Fristrom JW: Interactions and developmental effects of mutations in the Broad-Complex of Drosophila melanogaster. Genetics 1988, 118:247-259.
- Fletcher JC, Burtis KC, Hogness DS, Thummel CS: The Drosophila E74 gene is required for metamorphosis and plays a role in the polytene chromosome puffing response to ecdysone. Development 1995, 121:1455-1465.
- Andres AJ, Thummel CS: The Drosophila 63F early puff contains E63-1, an ecdysone-inducible gene that encodes a novel Ca²⁺binding protein. Development 1995, 121:2667-2679.
- Hock T, Cottrill T, Keegan J, Garza D: The E23 early gene of Drosophila encodes an ecdysone-inducible ATP-binding cassette transporter capable of repressing ecdysone-mediated gene activation. Proc Natl Acad Sci USA 2000, 97:9519-9524.
- Hurban P, Thummel CS: Isolation and characterization of fifteen ecdysone-inducible Drosophila genes reveal unexpected complexities in ecdysone regulation. Mol Cell Biol 1993, 13:7101-7111.
- Stowers RS, Garza D, Rascle A, Hogness DS: The L63 gene is necessary for the ecdysone-induced 63E late puff and encodes

CDK proteins required for Drosophila development. Dev Biol 2000, 221:23-40.

- Arbeitman MN, Furlong EE, Imam F, Johnson E, Null BH, Baker BS, Krasnow MA, Scott MP, Davis RW, White KP: Gene expression during the life cycle of Drosophila melanogaster. Science 2002, 297:2270-2275.
- 20. Li TR, White KP: Tissue-specific gene expression and ecdysone-regulated genomic networks in Drosophila. Dev Cell 2003, 5:59-72.
- 21. White KP, Rifkin SA, Hurban P, Hogness DS: Microarray analysis of Drosophila development during metamorphosis. Science 1999, 286:2179-2184.
- Baker KD, Shewchuk LM, Kozlova T, Makishima M, Hassell A, Wisely B, Caravella JA, Lambert MH, Reinking JL, Krause H, et al.: The Drosophila orphan nuclear receptor DHR38 mediates an atypical ecdysteroid signaling pathway. Cell 2003, 113:731-742.
- 23. Champlin DT, Truman JW: Ecdysteroid control of cell proliferation during optic lobe neurogenesis in the moth Manduca sexta. Development 1998, 125:269-277.
- 24. Bashirullah A, Pasquinelli AE, Kiger AA, Perrimon N, Ruvkun G, Thummel CS: Coordinate regulation of small temporal RNAs at the onset of Drosophila metamorphosis. Dev Biol 2003, 259:1-8.
- 25. Arama E, Dickman D, Kimchie Z, Shearn A, Lev Z: Mutations in the beta-propeller domain of the Drosophila brain tumor (brat) protein induce neoplasm in the larval brain. Oncogene 2000, 19:3706-3716.
- 26. Sonoda J, Wharton RP: Drosophila Brain Tumor is a translational repressor. Genes Dev 2001, 15:762-773.
- 27. Andres AJ, Fletcher JC, Karim FD, Thummel CS: Molecular analysis of the initiation of insect metamorphosis: a comparative study of Drosophila ecdysteroid-regulated transcription. Dev Biol 1993, 160:388-404.
- 28. Horner MA, Chen T, Thummel CS: Ecdysteroid regulation and DNA binding properties of Drosophila nuclear hormone receptor superfamily members. Dev Biol 1995, 168:490-502.
- Lam G, Thummel CS: Inducible expression of double-stranded RNA directs specific genetic interference in Drosophila. Curr Biol 2000, 10:957-963.
- Schubiger M, Truman JW: The RXR ortholog USP suppresses early metamorphic processes in *Drosophila* in the absence of ecdysteroids. *Development* 2000, 127:1151-1159.
 Clever U, Clever I, Storbeck I, Young NL: The apparent
- Clever U, Clever I, Storbeck I, Young NL: The apparent requirement of two hormones, α- and β-ecdysone, for molting induction in insects. Dev Biol 1973, 31:47-60.
- Oberlander H: α-ecdysone induced DNA synthesis in cultured wing discs of Galleria mellonella - Inhibition by 20-hydroxyecdysone and 22-isoecdysone. J Insect Phys 1972, 18:223-228.
- Hiruma K, Shinoda T, Malone F, Riddiford LM: Juvenile hormone modulates 20-hydroxyecdysone-inducible ecdysone receptor and ultraspiracle gene expression in the tobacco hornworm, Manduca sexta. Dev Genes Evol 1999, 209:18-30.
- Zhou B, Hiruma K, Jindra M, Shinoda T, Segraves WA, Malone F, Riddiford LM: Regulation of the transcription factor E75 by 20hydroxyecdysone and juvenile hormone in the epidermis of the tobacco hornworm, *Manduca sexta*, during larval molting and metamorphosis. Dev Biol 1998, 193:127-138.
- 35. Zhou B, Hiruma K, Shinoda T, Riddiford LM: Juvenile hormone prevents ecdysteroid-induced expression of broad complex RNAs in the epidermis of the tobacco hornworm, *Manduca* sexta. Dev Biol 1998, 203:233-244.
- 36. Richards G, Da Lage JL, Huet F, Ruiz C: The acquisition of competence to respond to ecdysone in Drosophila is transcript specific. Mech Dev 1999, 82:131-139.
- Lee CY, Clough EA, Yellon P, Teslovich TM, Stephan DA, Baehrecke EH: Genome-wide analyses of steroid- and radiation-triggered programmed cell death in Drosophila. Curr Biol 2003, 13:350-357.
- Pecasse F, Beck Y, Ruiz C, Richards G: Kruppel-homolog, a stagespecific modulator of the prepupal ecdysone response, is essential for Drosophila metamorphosis. Dev Biol 2000, 221:53-67.
- Koelle MR, Segraves WA, Hogness DS: DHR3: a Drosophila steroid receptor homolog. Proc Natl Acad Sci USA 1992, 89:6167-6171.
- Drysdale RA, Crosby MA, FlyBase Consortium: FlyBase: genes and gene models. Nucleic Acids Res 2005, 33:D390-D395.
- 41. McDonald MJ, Rosbash M: Microarray analysis and organization of circadian gene expression in Drosophila. Cell 2001,

107:567-578.

- Claridge-Chang A, Wijnen H, Naef F, Boothroyd C, Rajewsky N, Young MW: Circadian regulation of gene expression systems in the Drosophila head. Neuron 2001, 32:657-671.
- Girardot F, Monnier V, Tricoire H: Genome wide analysis of common and specific stress responses in adult Drosophila melanogaster. BMC Genomics 2004, 5:74.
- De Gregorio E, Spellman PT, Rubin GM, Lemaitre B: Genome-wide analysis of the Drosophila immune response by using oligonucleotide microarrays. Proc Natl Acad Sci USA 2001, 98:12590-12595.
- 45. Zinke I, Schutz CS, Katzenberger JD, Bauer M, Pankratz MJ: Nutrient control of gene expression in Drosophila: microarray analysis of starvation and sugar-dependent response. *EMBO J* 2002, 21:6162-6173.
- 46. Carlson JR, Hogness DS: Developmental and functional analysis of Jonah gene expression. Dev Biol 1985, 108:355-368.
- King-Jones K, Charles JP, Lam G, Thummel CS: The ecdysoneinduced DHR4 orphan nuclear receptor coordinates growth and maturation in Drosophila. Cell 2005, 121:773-784.
- Weber AN, Tauszig-Delamasure S, Hoffmann JA, Lelievre E, Gascan H, Ray KP, Morse MA, Imler JL, Gay NJ: Binding of the Drosophila cytokine Spatzle to Toll is direct and establishes signaling. Nat Immunol 2003, 4:794-800.
- Karim FD, Thummel CS: Ecdysone coordinates the timing and amounts of E74A and E74B transcription in Drosophila. Genes Dev 1991, 5:1067-1079.
- Loop T, Leemans R, Stiefel U, Hermida L, Egger B, Xie F, Primig M, Certa U, Fischbach KF, Reichert H, Hirth F: Transcriptional signature of an adult brain tumor in Drosophila. BMC Genomics 2004, 5:24.
- Bayer CA, von Kalm L, Fristrom JW: Relationships between protein isoforms and genetic functions demonstrate functional redundancy at the Broad-Complex during Drosophila metamorphosis. Dev Biol 1997, 187:267-282.
- Yamamoto KR, Alberts BM: Steroid receptors: elements for modulation of eukaryotic transcription. Annu Rev Biochem 1976, 45:721-746.
- Andres AJ, Thummel CS: Methods for quantitative analysis of transcription in larvae and prepupae. In Drosophila melanogaster: Practical Uses in Cell and Molecular Biology Edited by: Goldstein LSB, Fyrberg EA. New York: Academic Press; 1994:565-573.
 Li C, Wong WH: Model-based analysis of oligonucleotide
- Li C, Wong WH: Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. Proc Natl Acad Sci USA 2001, 98:31-36.
- Tusher VG, Tibshirani R, Chu G: Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci USA 2001, 98:5116-5121.
- 56. Gene Expression Omnibus [http://www.ncbi.nlm.nih.gov/geo/]