Professor of Biology:
Bruce A. Hay

Research Fellows:
Omar Akbari, Nikolai Kandul, Philippos Papathanos

Graduate Students:
Anna Buchman, Kelly J. Dusinberre, Jeremy Sandler

Undergraduate Students:
Ran Yang

Research Staff:
Haixia Huang, Katie Kennedy

Collaborators:
H.-A.J. Müller¹, M. Guo², Rollie Clem³, Yigong Shi⁴, Chun-Hong Chen⁵, S.J. Yoo⁶, Anthony James⁷, Zhijian Tu⁸, John M. Marshall⁹, Igor Antoshechkin¹⁰

¹Heinrich-Heine Universitat, Düsseldorf, Germany
²Department of Neurology, UCLA
³Kansas State University, Kansas
⁴Princeton University, New Jersey
⁵National Health Research Institutes, Taiwan
⁶Kyung Hee University, Seoul, Korea
⁷UC Irvine
⁸Virginia Tech
⁹Imperial College
¹⁰Caltech Genomics Facility

Financial Support:
Edelman Fund
Foundation for NIH Research
Hicks Alzheimer's Fund
National Institutes of Health
Weston Havens Foundation

**Summary:**
We are interested in multiple questions in basic and applied biology. For further information on Hay lab research consult our web page (http://www.its.caltech.edu/~haylab/). One goal of our work is directed towards understanding the genetic and molecular mechanisms that regulate cell
death, proliferation, innate immunity, microRNA function, and spermatogenesis. We use Drosophila melanogaster as a model system to identify genes that function to regulate these processes. Important cellular regulatory pathways are evolutionarily conserved; thus, molecules identified as regulators of these processes in Drosophila are likely to have homologs in vertebrates and the pathways that link these molecules are likely to be regulated similarly.

A second goal of our work addresses three questions in population biology. 1) Can we bring about reproductive isolation (speciation) between populations of plants or animals that otherwise freely interbreed? Answers to this question have application to the growing number of situations in which plants and animals are engineered to show specific pharmaceutical or agricultural traits. In brief, we would like to be able to limit gene flow between engineered organisms and their wild counterparts. 2) Can we engineer the genetics of populations so that they drive themselves to local extinction? For example, invasive non-native plants and animals cause substantial economic losses. A number also cause substantial environmental damage, leading in many cases to extensive range reduction and/or extinction of unique, endemic species. Our goal is to develop genetic tricks that drive local extinction of invasive species and disease vectors. 3) Can we drive genes into wild populations so that all individuals express a trait of interest? With regard to this last aim, we are particularly interested in developing transgenic insects that will prevent transmission of mosquito-borne pathogens that cause malaria and dengue fever. More than 500 million people are infected with the malaria parasite each year, resulting in 1-3 million deaths, while dengue, a mosquito-borne virus, infects more than 100 million people each year, resulting in more than 25,000 deaths. Effective vaccines do not exist, and in the case of malaria, the causative agent, the parasite Plasmodium falciparum has acquired resistance to many drugs. Vector suppression through the release of sterile males, the use of insecticides, or modification of the environment provides an important tool for limiting mosquito-borne disease. However, each approach has limitations. Release of sterile males provides only transient population suppression, insecticides affect many non-target species and mosquitoes often evolve resistance to these compounds, and wholesale modification of the environment may not be feasible, or desirable in many situations based on ecological concerns. Our goals are two-fold: to develop transgenic insects that lack the ability to transmit these pathogens (primarily as collaborations with other labs); and to develop genetic tools for driving these genes into wild populations of insects, thereby blocking disease transmission.

Approaches similar to those described above can also be used to tackle diseases of agricultural interest. One disease of current interest is known as citrus greening disease (also known as Huanglongbing; HLB). HLB is caused by the bacteria Candidatus Liberibacter, which is transmitted to the citrus plant by an insect, the phloem feeding citrus psyllid, Diaphorina citri. The disease is difficult to detect and current methods of control involve either regular use of insecticides or –once the tree is infected – tree destruction. HLB threatens to effectively eliminate the citrus industry in many areas in the US. We are interested in working with the citrus industry to develop transgenic insect-based approaches to prevent HLB.
We are also interested in the molecular mechanisms that underlie parthenogenesis – the ability of females to reproduce without males. While parthenogenesis is known to occur in many different animals, the molecular mechanisms that allow a switch from sexual to parthenogenetic reproduction are unknown for any species. To get at this question we have sequenced genomes of sexual and parthenogenetic versions of *Drosophila pallidosa* and are using classical mapping and RNA-seq approaches to identify the genes involved.

*Drosophila models of human neuro-degenerative diseases* (Ming Guo (and the Guo lab), Haixia Huang, Bruce A. Hay, Nikolai Kandul). In collaboration with the Guo lab at UCLA we are studying *Drosophila* models of the two most common neurodegenerative diseases, Alzheimer's disease and Parkinson's disease (Guo, M. et al. (2003) *Hum. Mol. Genet.* 12:2669-2678; Clark, I.E. et al. (2006) *Nature* 441:1162-1166). We are particularly interested in understanding how disruption of mitochondrial function contributes to these diseases.

**Gene activation screens for cell death regulators:** MicroRNAs, small non-coding RNAs, define a new family of cell death regulator (Haixia Huang, Bruce Hay). We have carried out several screens for cell death regulators in the fly and have identified a number of new molecules. Among these are multiple microRNAs, small noncoding RNAs that function by inhibiting translation of target transcripts. We are interested in determining when and where these molecules regulate death, as well as the nature of their targets. We are also designing microRNAs that target known cell death regulators as a way of probing the function of these proteins in specific contexts.

**Cell death, caspases and IAPs** (*H. Arno J. Müller, Soon Ji Yoo, Bruce A. Hay*). In flies and vertebrates most, if not all, cells can undergo apoptosis in the absence of new gene expression, indicating that the components required to carry out apoptosis are present and ready for activation. The core of the cell death machine consists of members of a family of proteases known as caspases, which become activated in response to many different death signals. Active caspases then cleave a number of different cellular substrates that ultimately lead to cell death and corpse phagocytosis. Most if not all cells constitutively express caspase zymogens (inactive precursors) sufficient to bring about apoptosis. Thus, the key to cell death and survival signaling revolves around controlling the levels of active caspases in the cell. Several basic strategies are used to regulate caspase activity, and the core proteins that drive caspase-dependent death are evolutionarily conserved. In *Drosophila* many cells experience chronic activation of the apical cell death caspase Dronc. If unrestrained, active Dronc cleaves and activates downstream effector caspases that bring about cell death. Cells survive because they express the IAP DIAP1, which suppresses Dronc activity, as well as that of caspases activated by Dronc. One major pathway through which caspase-dependent cell death in flies is induced is through the regulated expression of pro-apoptotic proteins that disrupt DIAP1-caspase interactions through several different mechanisms, each of which has the effect of unleashing a cascade of apoptosis-inducing caspase activity. We are interested in several questions. 1) What are the signals that lead to caspase activation in cells that would normally live? 2) How do IAPs regulate caspase activity and when and where does this regulation define points of control? 3) How is IAP activity regulated? 4) And
finally, as discussed further below, how do caspases, IAPs and their regulators work to regulate non-apoptotic processes? We are using both genetic screens and biochemical approaches to identify the critical molecules.

**Caspases and their regulators in a non-apoptotic process, spermatid differentiation** (*Haixia Huang, Geoffrey Pittman*). We have found that multiple caspases, acting through distinct pathways, acting at distinct points in time and space, are required for spermatid individualization, a process in which spermatids (which develop in a common cytoplasm) become enclosed in individual plasma membranes and shed most of their cytoplasm Huh, J. *et al.* (2004) *PLoS Biology* 1:E15. Spermatid individualization is an evolutionarily conserved process, but little is known about how it is brought about. Several questions are of interest to us: 1) What are the upstream signals that drive caspase activation? 2) What are the nonapoptotic targets that facilitate differentiation? 3) How is cell death prevented in the face of high levels of caspase activity that would normally be associated with cell death? 4) Do caspases play similar roles in promoting spermatid differentiation in mammals? 5) Can we manipulate the biology of spermatogenesis so as to bias gamete production so that males produce gametes carrying the Y chromosome, but not the X chromosome? Elements with these characteristics, if they are located on the Y chromosome, are predicted to drive a population to extinction through the generation of male-only populations.

**Cell death and the innate immune system** (*Bruce A. Hay*). As discussed above, many IAP family proteins inhibit apoptosis. IAPs contain N-terminal BIR domains and a C-terminal RING ubiquitin ligase domain. *Drosophila* DIAP1 protects cells from apoptosis by inhibiting caspases. Apoptosis initiates when proteins such as Reaper and Hid bind a surface groove in DIAP1 BIR domains via an N-terminal IAP-binding motif (IBM). This evolutionarily conserved interaction disrupts IAP-caspase interactions, unleashing apoptosis-inducing caspase activity. DIAP2 overexpression also inhibits Rpr- and Hid-dependent apoptosis, but little is known about DIAP2's normal functions. We generated *diap2* null mutants, which are viable and show no defects in developmental or stress-induced apoptosis. Instead, DIAP2 is required for the innate immune response to Gram-negative bacterial infection (Huh, J. *et al.* (2007) *J. Biol. Chem.* 282:2056-2068). DIAP2 promotes cytoplasmic cleavage and nuclear translocation of the NF-kB homolog Relish, and this requires the DIAP2 RING domain. Increasing the genetic dose of *diap2* results in an increased immune response, while expression of Rpr or Hid results in down-regulation of DIAP2 protein levels. Together these observations suggest that DIAP2 can regulate immune signaling in a dose-dependent manner, and that DIAP2 is regulated by IBM-containing proteins. Therefore, *diap2* may identify a point of convergence between apoptosis and immune signaling pathways.

**Driving genes for disease refractoriness into wild pest insect populations with Medea selfish genetic elements** (*Haixia Huang, Catherine Ward, Geoff Pittman, Omar Akbari, Arun Kumar, Zachary Sun, Philippos Papathanos, Jeremy Sandler, Bruce A. Hay*). An attractive approach to suppressing mosquito-borne diseases involves replacing the wild-insect population with modified counterparts unable to transmit disease. Mosquitoes with a diminished capacity to transmit *Plasmodium* have been identified in the wild and created in the laboratory, demonstrating that
endogenous or engineered mosquito immunity can be harnessed to attack *Plasmodium*. However, a critical unanswered question is how to spread these effector genes throughout the areas inhabited by disease-transmitting insects. Epidemiological and modeling studies suggest that it will be necessary to rapidly replace a large percentage of the wild mosquito population with refractory insects in order to achieve significant levels of disease control. Because insect disease vectors are spread over wide areas and can migrate significant distances, mass release of refractory insects associated with simple Mendelian transmission of effector-bearing chromosomes is unlikely to result in a high enough frequency of transgene-bearing individuals. Compounding this problem, enhancement of immune function in insects is often costly, requiring tradeoffs with other life history traits such as longevity and fecundity that decrease fitness. Therefore, it is likely that insects carrying effector transgenes will be less fit than their wild counterparts, resulting in a decrease in the fraction of individuals carrying genes for refractoriness over time. These observations argue that population replacement will require coupling of genes conferring disease refractoriness with a genetic mechanism for driving these genes through the wild population at greater than Mendelian frequencies.

Maternal-effect lethal selfish genetic elements in the flour beetle *Tribolium castaneum* have the following behavior: when present in a female, they must be inherited in the next generation in order for the offspring to survive. The molecular nature of these elements (known as Medea elements) is unknown, but their spiteful genetic behavior (they cause the death of those who fail to inherit them, giving a relative transmission advantage to those that do carry them) makes them attractive candidates to mediate drive because it is predicted to lead to rapid spread of the element within the population even if it carries an associated fitness cost. Medea's ability to spread, and the time it takes to become present in all individuals, is a function of fitness cost and introduction frequency. The plot in Figure 1 describes the number of generations required for *Medea* to be present in 99% of individuals, for a *Medea* element with an embryonic fitness cost (resulting from the presence of a cargo transgene designed to protect from disease, for example). Homozygous *Medea*:non-*Medea* introduction ratios are indicated on the Y axis, and embryonic fitness cost on the X axis. Area between lines indicates regions of parameter space within which a specific number of generations (indicated by numbers and arrows) are required for the frequency of *Medea* individuals to reach a frequency of 99% or greater. Line color, shown in the heat map at right, provides a measure of how many generations are required. Black lines (50+) indicate that fifty or more generations are required. The border between the black-lined region and
the lower unlined region defines the critical Medea:non-Medea introduction ratio, below which Medea will be eliminated from the population.

The molecular biology of endogenous Medea elements is unknown, but the genetics suggests a model in which Medea consists of two linked genes: The first encodes a toxin that is expressed only in the female germline, with effects that are passed to all progeny. The second encodes an antidote, expressed under the control of an early zygote-specific promoter (Figure 2).

Mothers that carry a Medea element express a toxin (red dots) that is inherited by all oocytes (small ovals). Embryos (large ovals) that do not inherit Medea die because toxin activity (red background) is unimpeded (lower left square). Embryos that inherit Medea from the mother (upper left square), the father (lower right square) or both (upper right square), survive because expression of an antidote early during embryogenesis (green background) neutralizes toxin activity. We imagine that Medea is comprised of two closely linked genes (upper left).

We created synthetic Medea elements in Drosophila that can drive population replacement (Figure 4) and that are resistant to recombination-mediated dissociation of drive and effector functions. These elements (Figure 3) result from zygotic rescue of a maternal loss-of-function that results in embryonic arrest. During oogenesis a maternal transcript is synthesized (green dots), whose product is required for early embryogenesis. In females carrying a Medea, the first transgene (the toxin) drives maternal drives maternal germline-specific expression of microRNAs that silence expression of the gene whose product is required for early embryogenesis. This results in
inheritance of a lethal condition - the loss of an essential maternally deposited product - by all oocytes/embryos. Progeny survive the embryonic arrest thereby induced if they inherit from their mother a tightly linked transgene driving early zygotic expression of the maternally silenced gene just in time to restore embryo development, but they die if they fail to inherit it.

**Sensing and killing dengue and yellow fever virus-infected cells in their insect host** *(Kelly J. Dusinberre, Zachary Sun)* Dengue and yellow Fever virus infect mosquitoes during a blood meal. The virus must enter and replicate inside mosquito midgut cells, disseminate throughout the body and ultimately infect the salivary gland (7-14 days later), in order to be transmitted to a new individual during a subsequent blood meal. Our goal is to develop transgenes that are phenotypically neutral when expressed in uninfected individuals, but that kill virus-infected cells and/or the mosquitoes themselves. The virus encodes several activities that are not present in uninfected host cells. These include a viral polyprotein protease, and RNA-dependent RNA polymerase. We are developing molecules that sense these activities and cause the death of cells and insects in which they occur, thereby preventing disease transmission to humans.

**Figure 5**

![Diagram showing genetic interactions](image)

**Engineering reproductive isolation and population replacement using a synthetic underdominance system** *(Kelly Dusinberre, Katie Kennedy, Mario Zuba, Jennifer Hu, Anna Buchman)*. The Medea system detailed above is very good at spreading genes into populations distributed over large areas, provided that modest levels of migration occur. This is ideal for situations in which the goal is to carry out population replacement in large regions. However, some communities may favor an approach in which population replacement is restricted to a local environment (Let's see how it does in your back yard, before trying it in mine). This creates a challenge: how to spread genes within a local environment, but maintain a barrier to migration-driven spread and fixation in surrounding regions. To address this need we are developing the
synthetic underdominance system illustrated in Figure 5. In this system homologous chromosomes carry toxin-antidote pairs in which the toxin present on chromosome A (Killer 1) is linked to an antidote (Rescue 2) that represses Killer 2. Killer 2 is located at the same position on the homologous chromosome B, linked with an antidote (Rescue 1) that represses Killer 1 (Figure x). In such a system, organisms can only survive if they carry A and B chromosomes (in A/B individuals), or only wildtype (+) chromosomes (in +/+ individuals). A/+ and B/+ individuals die. A and B chromosomes will also carry genes that confer resistance to disease transmission. Such a system has two interesting features.

First, it constitutes a simple method for engineering reproductive isolation (speciation). Matings between +/+ individuals produce viable progeny, as do matings between A/B individuals. However, mating between +/+ and A/B individuals produce only A/+ and B/+ progeny, which all die. This simple technology has a number of potential applications and provides a platform from which to explore some of the evolutionary consequences of reproductive isolation. Second, it provides a method for driving genes into a local environment in such a way that they are unlikely spread to fixation in surrounding regions through migration. In brief, for underdominance, as with Medea elements that carry a fitness cost, a threshold frequency must be achieved in order for spread to occur at all. With single locus underdominance this threshold is quite high (66%) (Figure 6, left panel). In two-locus underdominance (Figure 6, right panel), the two toxin-antidote cassettes are located on non-homologous chromosomes. In this configuration more transgenic progeny can survive in crosses to wildtype, and thus the introduction threshold required for spread to occur is significantly lower, 33%. Once the threshold is crossed, these underdominant systems drive the wildtype chromosomes out of the population by causing their death in individuals that carry A or B, but not both. The A/B genotypes have great difficulty in spreading into surrounding regions through migration because as they migrate into areas composed largely of +/+ individuals, they are more likely to mate with +/+ individuals than with A/B individuals, resulting in the likely death of progeny that carry one but not the other. We are
developing several versions of underdominance in Drosophila and are working to move these systems to mosquito species.

**Sensing and responding to normal and abnormal microRNA expression** (Nikolai Kandul). MicroRNAs (miRNAs) are small, non-coding RNAs that regulate gene expression by suppressing the translation or promoting the degradation of transcripts to which they hybridize. Importantly for our purposes, when miRNAs are perfectly complementary to their target transcripts, transcript cleavage and degradation results. It is clear that miRNA expression is deregulated in many disease states. In addition, many viruses encode miRNAs that promote viral replication and/or suppress host defense systems. Our goal is to develop methods for sensing the expression of a particular miRNA, and then transducing this signal into changes in gene or protein expression. This will allow us to monitor the levels of miRNA expression in living animals. It will also allow us to regulate cellular physiology in response to the levels of particular miRNAs.

**Predicting the fate of gene drive systems and their cargos in the wild** (John Marshall, Bruce Hay, Catherine Ward, Jennifer Hu). As we develop gene drive strategies we need to be able to predict how they are likely to behave. A number of questions arise: Under what ecological and population genetic conditions will drive chromosomes spread? What are the likely epidemiological consequences of spread in terms of disease prevention? What are the likely functional lifetimes of these elements in the wild? What are the possibilities for removal and replacement of first-generation elements with second-generation elements? We are using mathematical modeling and computer simulations to address these issues for a number of different drive strategies.

**How many possible ways are there for driving genes into populations, resulting in either population replacement or population elimination?** (John Marshall, Bruce Hay). We are interested in identifying all the ways in which genes, gene complexes, or entire chromosomes can promote their own spread into populations. This analysis may identify novel mechanisms by which populations have been shaped in the wild. It may also identify mechanisms that could be used to drive genes into populations, either providing them with some desirable trait, or driving the population towards an inviable genotype and extinction. We are particularly interested in identifying those mechanisms that can be thought of as consisting of combinations of genes with toxin and antidote activities as these can in principal be engineered, and may also have evolved in the wild as a consequence of epistatic interactions between genes.
2012


2011


