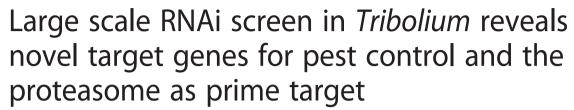


### **RESEARCH ARTICLE**

**Open Access** 





Julia Ulrich<sup>1</sup>, Van Anh Dao<sup>3</sup>, Upalparna Majumdar<sup>2</sup>, Christian Schmitt-Engel<sup>1</sup>, Jonas Schwirz<sup>1</sup>, Dorothea Schultheis<sup>2</sup>, Nadi Ströhlein<sup>2</sup>, Nicole Troelenberg<sup>2</sup>, Daniela Grossmann<sup>1</sup>, Tobias Richter<sup>2</sup>, Jürgen Dönitz<sup>1</sup>, Lizzy Gerischer<sup>4</sup>, Gérard Leboulle<sup>5</sup>, Andreas Vilcinskas<sup>6</sup>, Mario Stanke<sup>4</sup> and Gregor Bucher<sup>1\*</sup>

#### Abstract

**Background:** Insect pest control is challenged by insecticide resistance and negative impact on ecology and health. One promising pest specific alternative is the generation of transgenic plants, which express double stranded RNAs targeting essential genes of a pest species. Upon feeding, the dsRNA induces gene silencing in the pest resulting in its death. However, the identification of efficient RNAi target genes remains a major challenge as genomic tools and breeding capacity is limited in most pest insects impeding whole-animal-high-throughput-screening.

**Results:** We use the red flour beetle *Tribolium castaneum* as a screening platform in order to identify the most efficient RNAi target genes. From about 5,000 randomly screened genes of the iBeetle RNAi screen we identify 11 novel and highly efficient RNAi targets. Our data allowed us to determine GO term combinations that are predictive for efficient RNAi target genes with proteasomal genes being most predictive. Finally, we show that RNAi target genes do not appear to act synergistically and that protein sequence conservation does not correlate with the number of potential off target sites.

**Conclusions:** Our results will aid the identification of RNAi target genes in many pest species by providing a manageable number of excellent candidate genes to be tested and the proteasome as prime target. Further, the identified GO term combinations will help to identify efficient target genes from organ specific transcriptomes. Our off target analysis is relevant for the sequence selection used in transgenic plants.

**Keywords:** Pest control, RNAi, RNAi target gene, iBeetle, Tribolium, Proteasome

### **Background**

Alternative strategies for sustainable agriculture attract increasing attention because cultural practices and agrochemicals are not sufficient or permanently successful in keeping plant pathogens and pest insects under control. One of the most challenging problems in modern plant protections is the rapid adaptation of pest insects to insecticides which imposes both increasing costs of chemical plant protection and growing public concern about the hazards on environment and human health [1–3]. Transgenic plants expressing *Bacillus thuringiensis* (Bt) toxins

provide an alternative option to engineer insect-resistant crops. Despite the still lasting public debate about the introduction of genetically modified plants into the field, the production of transgenic plants increases globally. However, economic important pest insects such as the western corn rootworm, *Diabrothica virgifera virgifera*, evolved resistance against Bt-toxins expressed in engineered maize [4]. The questionable sustainability of this "first generation transgenic crops" and the fact that Bt-toxins can also impair non-target insects raised a demand for alternative and more pest specific transgenic approaches in plant protection among which the RNA-Interference (RNAi) appears most promising [5, 6]. Targeted silencing of gene expression by injection of corresponding double-stranded RNA (dsRNA) has become a major tool in functional analysis of genes in

<sup>&</sup>lt;sup>1</sup>Johann-Friedrich-Blumenbach-Institut, GZMB, Georg-August-Universität Göttingen, Justus-von-Liebig-Weg 11, 37077 Göttingen, Germany Full list of author information is available at the end of the article



<sup>\*</sup> Correspondence: gbucher1@uni-goettingen.de

insects [7-10]. In many insects, like the red flour beetle Tribolium castaneum, the RNAi effect spreads throughout the animal and is even transmitted to the offspring [11–14]. Importantly, some insects elicit the RNAi response upon ingestion of dsRNA. Based on this, dsRNAs have been implicated as a new generation of speciesspecific insecticides that kill insects upon oral up-take [15, 16]. Consequently, the possibility to engineer crops expressing double-stranded RNA which can silence target genes in pest insects after ingestion provides the opportunity for the development of species-specific control measures. One current limitation is the identification of dsRNAs that kill the targeted pest insect even when provided in minute amounts. First, the production of dsRNA in transgenic crops is limited and second, only small amounts are delivered to the pest insect via the ingested plant food. Unfortunately, the systematic identification of suitable RNAi target genes is challenging in pest species because they are usually difficult to rear in the lab and often lack the required genomic and genetic tools for a whole-animal-high-throughput-screen. The main insect model system Drosophila melanogaster lacks systemic RNAi and is therefore not well suited to screen for RNAi target genes [17].

The red flour beetle *Tribolium castaneum* has developed to an excellent insect model organism. It can be reared in large amounts in the lab, reproduces all year round, was the first beetle to be sequenced and both genetic and transgenic tools are available [18–22]. Its main strength is the strong and systemic RNAi response [10–12], which allowed performing a large scale unbiased RNAi screen (iBeetle Screen) [23, 24].

The currently prevailing small scale approaches of RNAi target identification may have missed the most efficient RNAi target genes. We reasoned that the red flour beetle would be an appropriate screening platform for the independent large scale identification of the most efficient RNAi target genes in insects. The respective orthologs could then be tested as RNAi target genes in other pest species. Therefore, we mined approximately 5.000 experiments of the iBeetle screen in order to identify those treatments that induced death of the injected animals most rapidly. We tested selected dsRNAs further by titration experiments and found that a number of them performed better than previously used RNAi target genes. Hence, the orthologs of these genes are novel prime candidates for RNAi based approaches of pest control of other pest species. Based on this set of efficient RNAi target genes, we identified GO term combinations that are predictive of good RNAi target genes and which identify the proteasome as a prime target. Finally, we tested and refuted the hypothesis that the efficiency of RNAi mediated pest control could be enhanced by synergistic action in double knock-downs. Further, we find that potential off-target-sequences occur independently of protein sequence conservation. Hence, off target effects can hardly be ruled out completely and efforts to reduce ecological side effects need to focus on selected species to be protected.

### Results and discussion

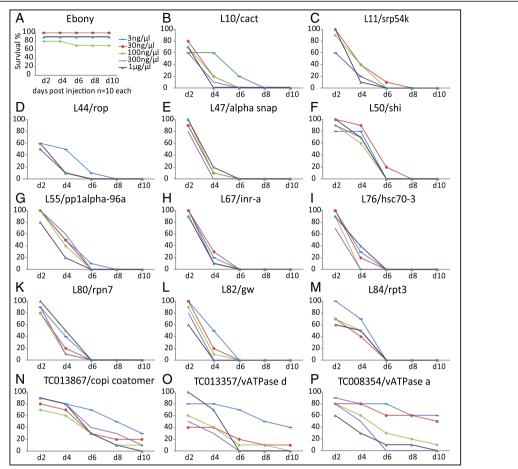
# Large scale RNAi screen identifies novel RNAi target genes

So far, the targets for dsRNA based pest control have been identified by small scale screens and on knowledge based approaches, i.e. by testing genes where previous data indicated an essential function. However, this approach will miss many genes that have not yet been linked to an essential function in one of the model species. Therefore, we screened the data produced by the large scale RNAi screen iBeetle (Bucher, Klingler, personal communication), where randomly selected genes were downregulated by injection into pupae and larvae and the resulting phenotypes were documented in the iBeetle-Base [23]. In the iBeetle screen, about 5,000 genes had been screened [24]. Of those, 100 revealed ≥90 % mortality both 9 days after pupal and eleven days after larval dsRNA injection (Additional file 1: Table S1). In order to confirm these results and to test for sensitivity, we injected different concentrations (3 ng/µl, 30 ng/ μl, 100 ng/μl, 300 ng/μl and 1 μg/μl) of the same dsRNAs into 10 larvae, respectively, and scored the survival rate every second day. The most effective 40 genes caused a mortality of 50-100 % at day eight post injection using the lowest concentrated dsRNA (Additional file 1: Table S2, Additional file 2: Figure S1 A). We focused on the 11 most effective target genes, which were marked by mortality of 100 % on day eight and at least 80 % on day six post injection (Fig. 1b-m). This high degree of lethality was confirmed by repeating the experiment using non-overlapping dsRNA fragments (1 µg/µl) making off target effects improbable (Additional file 1: Table S3).

For comparison, we performed the same experiment with the orthologs of five RNAi target genes published in the seminal paper of Baum et al. [15], which caused lethality in the western corn rootworm (WCR) upon dsRNA ingestion. Indeed, the *Tribolium* orthologs of these genes induced a high degree of mortality, but especially with low dsRNA concentrations, the mortality did not reach the one of the 11 candidates identified in this study (Fig. 1n–p, Additional file 2: Figure S1 B).

In order to check for stage dependence of the lethal effect, we injected dsRNAs targeting these 11 genes into ten adult beetles, respectively. Mortality rate on day eight post treatment was at least 90 % with 100 ng/ $\mu$ l dsRNA concentration while injection of 3 ng/ $\mu$ l dsRNA led to lower degree of mortality indicating a concentration response curve at these concentrations (Additional file 2: Figure S2). In summary, we identified 11 novel RNAi target genes that efficiently and rapidly induce lethality at larval, pupal and

Ulrich et al. BMC Genomics (2015) 16:674



**Fig. 1** Identification of novel RNAi target genes in *Tribolium castaneum*. The survival after knock-down of the 11 most efficient RNAi target genes identified in our screen is shown. dsRNAs of different concentrations were injected into 10 *Tribolium* larvae, respectively, and the survival rate was recorded every second day. **a** The pigmentation gene *Tc-ebony* was used as negative control. **b**–**m** Most of our novel RNAi target genes showed a larval mortality of 100 % on day eight and 80 % on day six post treatment for every dsRNA concentration. **n**–**p** RNAi treatment of commonly used targets based on the seminal paper of Baum et al. [15] revealed a lower efficiency (see also Additional file 2: Fig. S1B)

adult stages even at low doses of dsRNA and are more efficient than previously used target genes at least in *Tribolium*.

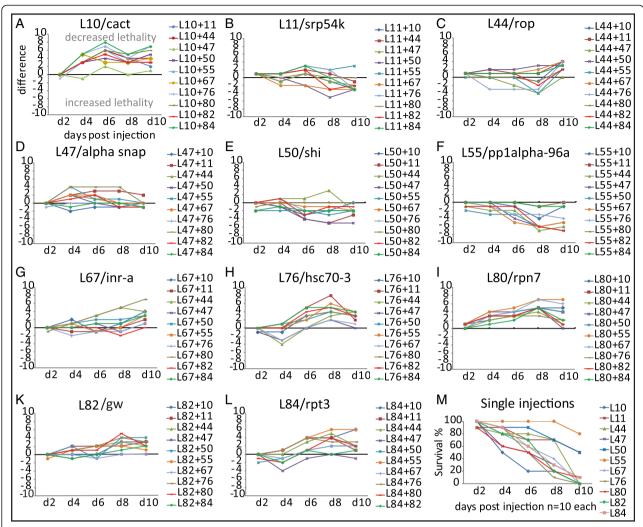
### Double RNAi led to additive but not to synergistic effects

We asked, whether the lethality of RNAi treatments can be increased synergistically by combined injection of two dsRNAs targeting different essential genes. All 55 pairwise combinations of the 11 top RNAi target genes were injected into larvae at the same end concentration as the single injections (i.e. either 0.5 ng/ $\mu$ l of one dsRNA or 0.25 ng/ $\mu$ l of two dsRNAs summing up to 0.5 ng/ $\mu$ l end concentration) and the survival was documented (Additional file 1: Table S4). We find no indication for synergism. Instead, the observed deviations from the baseline in some combinations are explained by additive effects: The most efficient targets become less penetrant when "diluted" with less effective dsRNAs (e.g. Fig. 2a) while less efficient targets become more potent when supplemented with

stronger target genes (e.g. Fig. 2f). In conclusion, we found no indication for synergistic effects that would be able to significantly enhance the technique (Fig. 2).

# Degree of sequence conservation does not strongly influence the number of off-targets

In order to protect non-target organisms it would be desirable to use dsRNA fragments that are specific to the pest species and do not contain sequences targeting genes in non-target organisms (off targets). Therefore, we asked whether protein sequence conservation of our RNAi target genes correlated with the number of potential off target sites in other species. On the protein sequence level most of our RNAi target genes showed a strong conservation between some well sequenced species covering insect diversity (*Drosophila melanogaster, Aedes aegypti* (Diptera), *Apis mellifera* (Hymenoptera), *Acyrthosiphon pisum* (Hemiptera). L10, L67 and L82 were the least conserved (Fig. 3a). For protein L76 no ortholog could be



**Fig. 2** Double RNAi led to additive but not synergistic effects. **a–I** Double RNAi treatments were performed and the number of surviving animals at different days post injection was documented (see values in Additional file 1: Table S4). From these values, the number of surviving animals of the respective single treatments (**m**) was subtracted, such that higher lethality in the double treatment is indicated by negative values in the panels. For every target gene, ten different pairwise combinations were injected with a total dsRNA concentration of 0.5 ng/µl and compared to the single injections (0.5 ng/µl total concentration). **m** Single injection of the 11 RNAi target genes were performed with a dsRNA concentration of 0.5 ng/µl

identified in *Aedes aegypti* (Fig. 3a; see Additional file 2: Figure S3 H for phylogenetic analysis).

DsRNAs are processed by the enzyme Dicer into 21-23 nt long short interfering RNAs (siRNAs). After incorporation into the RNA-induced silencing complex (RISC), they serve as template to recognize the complementary mRNA and target it for destruction [9]. However, siRNAs with an exact sequence identity of  $\geq 17$  nt can already induce off target effects [25]. Therefore, we searched the nucleotide sequences of the 11 *Tribolium* target genes against the well annotated NCBI transcriptome databases of the abovementioned species for  $\geq 17$  bp long stretches of identity (http://blast.ncbi.nlm.nih.gov/Blast.cgi) [26]. For visualization, we plotted these putative off target regions targeting in

genes of other species against the RNAi target gene nucleotide sequence (Fig. 4).

We did not find an overt correlation between the conservation of a protein and the number of potential off target regions (compare the most diverged genes L10, L67 and L82 with more conserved genes in Fig. 4). Likewise, within a given gene we found no enrichment of potential off target regions in more conserved stretches of the sequence (e.g. conserved protein domains) compared to less conserved stretches (e.g. non-coding UTRs; Fig. 4). Importantly, the location of potential off target sites was generally different for the different species. Together, these observations indicate that the number and location of the off target sites does not strongly correlate with protein sequence conservation. Further, given the vast diversity of taxa even in small

Α	Symbol	Tc-gene	Dm-or	tholog	Dm gene name		Gene symbol	Dm	Am	Aa	Ар	
	L10	TC002003	CG5848		cactus		cact	25.3%	32.9%	29.1%	20.4%	
	L11	TC002574	CG4659		signal recognition particle protein 54k		srp54k	88.5%	88.2%	73.8%	87.2%	
	L44	TC011120	CG15811		ras opposite		rop	76.7%	76.4%	75.9%	76.3%	
	L47	TC013571	CG6625		alpha soluble NSF attachment protein		alpha snap	75.3%	84.2%	82.3%	73.6%	
	L50	TC011058	CG18	3102	shibire		shi	79.6%	86.0%	80.7%	82.5%	
	L55	TC015321	CG65	593	protein phosphatase 1alpha at 96a		pp1alpha-96a	92.0%	94.8%	92.1%	92.1%	
	L67	TC008263	CG10	0228	inverse regulator a		inr-a	22.2%	33.9%	23.9%	25.0%	
	L76	TC004425	CG4:	147	heat shock 70-kDa protein cognate 3		hsc70-3	91.5%	90.4%	-	88.6%	
	L80	TC006375	CG53	378	regulatory particle non-ATPase 7		rpn7	72.0%	74.6%	72.8%	37,8%	
	L82	TC006679	CG3:	1992	gawky		gw	29.4%	39.3%	31.7%	36.1%	
	L84				regulatory particle triple-A ATPase 3		rpt3	91.7%	91.8%	92.9%	90.6%	
В	Lethality of novel genes		Cluster 1 Clus		ster 2		Cluster 3					
	with clustering of Top40						ber in % n=3)	(gene number in % n=69)				
	70-100% lethality on d11				86% (6 of 7)	100%	(3 of 3)	45% (31 of 69)				
					Cluster 5	Clus	ster 6	Cluster 7				
	(gene				ne number in % n=55)	(gene numb	er in % n=243)	(gene number in % n=66)				
	70-100%	lethality on	d11	1 24% (13 of 55)		32% (77 of 243)		56% (37 of 66)				
					Cluster 8 Cluster 10							
	(gene				ne number in % n=35)	(gene numb	er in % n=19)					
	70-100% lethality on d11				49% (17 of 35)	42% (8 of 19)						
С	Annotation Cluster 1 E		Enric	Enrichment Score: 4,26								
	Category Term			1						/ D\	, ,	
	SP_PIR_KEYWORDS proteasome			•					Count 9		/alue	
				easome					7	19 1	,61E-09	
	GOTERM_	_CC_FAT	GO:0	easome 000502	~proteasome complex				7 7	19 1 19 1	,61E-09 ,03E-06	
	KEGG_PA	_CC_FAT THWAY	GO:0 dme0	easome 000502 03050:P	roteasome				7 7 7	19 1 19 1 19 1	,61E-09 ,03E-06 ,37E-06	
	KEGG_PA	_CC_FAT THWAY _MF_FAT	GO:0 dme0 GO:0	easome 000502 03050:P 004175	roteasome rendopeptidase activity				7 7 7 7	19 1 19 1 19 1 19 1	,61E-09 ,03E-06 ,37E-06 ,21E-02	
	KEGG_PA GOTERM_ GOTERM	_CC_FAT THWAY _MF_FAT _BP_FAT	GO:0 dme0 GO:0 GO:0	easome 000502 03050:P 004175 006508	roteasome ~endopeptidase activity ~proteolysis	g on I-amino ao	id nentides		7 7 7 7 8	19 1 19 1 19 1 19 1 22 2	,61E-09 ,03E-06 ,37E-06 ,21E-02 ,39E-02	
	KEGG_PA GOTERM_ GOTERM_ GOTERM_	_CC_FAT THWAY _MF_FAT _BP_FAT	GO:0 dme0 GO:0 GO:0	easome 000502 03050:P 004175 006508 070011	roteasome ~endopeptidase activity ~proteolysis ~peptidase activity, actin	•			7 7 7 7	19 1 19 1 19 1 19 1 22 2	,61E-09 ,03E-06 ,37E-06 ,21E-02	
	KEGG_PA' GOTERM_ GOTERM_ GOTERM_ Genes	_CC_FAT THWAY _MF_FAT _BP_FAT	GO:0 dme0 GO:0 GO:0 GO:0	easome 000502 03050:P 004175 006508 070011	roteasome ~endopeptidase activity ~proteolysis	•			7 7 7 7 8	19 1 19 1 19 1 19 1 22 2	,61E-09 ,03E-06 ,37E-06 ,21E-02 ,39E-02	
	KEGG_PA' GOTERM GOTERM GOTERM Genes Annotation	CC_FAT THWAY _MF_FAT _BP_FAT _MF_FAT on Cluster 2	GO:0 dme0 GO:0 GO:0 GO:0 CG16	easome 000502 03050:P 004175 006508 070011 6916, CG hment \$	oteasome  rendopeptidase activity  roteolysis  reptidase activity, actin 5378, CG18174, CG4157 core: 2,37	•			7 7 7 7 8	19 1 19 1 19 1 19 1 22 2 19 4	,61E-09 ,03E-06 ,37E-06 ,21E-02 ,39E-02 ,36E-02	
	KEGG_PA' GOTERM GOTERM GOTERM Genes Annotation	CC_FAT THWAY MF_FAT BP_FAT MF_FAT on Cluster 2 EYWORDS	GO:0 dme0 GO:0 GO:0 GO:0 CG16 Enric	easome 000502 03050:P 004175 006508 070011 6916, CG hment 9	oteasome  rendopeptidase activity  roteolysis  reptidase activity, actin 5378, CG18174, CG4157 core: 2,37	, CG4904, CG17			7 7 7 7 8 7	19 1 19 1 19 1 19 1 22 2 19 4	,61E-09 ,03E-06 ,37E-06 ,21E-02 ,39E-02	
	KEGG_PA' GOTERM_ GOTERM_ GOTERM_ Genes Annotation SP_PIR_K	CC_FAT THWAY _MF_FAT _BP_FAT _MF_FAT on Cluster 2 EYWORDS	GO:0 dme0 GO:0 GO:0 GO:0 CG16 Enric	easome 000502 03050:P 004175 006508 070011 5916, CG hment S	oteasome  rendopeptidase activity  proteolysis  repetidase activity, activity  5378, CG18174, CG4157  core: 2,37	, CG4904, CG17 a/beta			7 7 7 7 8 7	19 1 19 1 19 1 19 1 22 2 19 4	,61E-09 ,03E-06 ,37E-06 ,21E-02 ,39E-02 ,36E-02	
	KEGG_PA GOTERM GOTERM GOTERM Genes Annotation SP_PIR_K INTERPRO	CC_FAT THWAY _MF_FAT _BP_FAT _MF_FAT on Cluster 2 EYWORDS ) _MF_FAT	GO:0 dmeC GO:0 GO:0 GO:0 CG16 Enric threc IPRO0 GO:0	easome 000502 03050:P 004175 006508 070011 6916, CG hment \$ onine pro 01353:P 004298	oteasome  "endopeptidase activity "proteolysis "peptidase activity, actin "5378, CG18174, CG4157 core: 2,37 otease roteasome, subunit alph	, CG4904, CG17 na/beta otidase activity			7 7 7 7 8 7	19 1 19 1 19 1 19 1 22 2 19 4	,61E-09 ,03E-06 ,37E-06 ,21E-02 ,39E-02 ,36E-02 ,72E-03 ,46E-03	
	KEGG_PA GOTERM GOTERM GOTERM Genes Annotati SP_PIR_K INTERPRO	CC_FAT THWAY _MF_FAT _BP_FAT _MF_FAT on Cluster 2 EYWORDS ) _MF_FAT _MF_FAT	GO:0 dmeC GO:0 GO:0 GO:0 CG16 Enric threc IPRO( GO:0 GO:0	easome 000502 03050:P 004175 006508 070011 6916, CG hment \$ online pro 01353:P 004298 070003	oteasome  rendopeptidase activity  roteolysis  peptidase activity, actin 5378, CG18174, CG4157  core: 2,37  otease  roteasome, subunit alph  rthreonine-type endoper	a/beta otidase activity			7 7 7 7 8 7	19 1 19 1 19 1 19 1 22 2 19 4 8 1 8 3 8 5 8 5	,61E-09 ,03E-06 ,37E-06 ,21E-02 ,39E-02 ,36E-02 ,72E-03 ,46E-03 ,05E-03	
	KEGG_PA GOTERM GOTERM GOTERM Genes Annotation SP_PIR_K INTERPRO GOTERM GOTERM	CC_FAT THWAY _MF_FAT _BP_FAT _MF_FAT on Cluster 2 EYWORDS ) _MF_FAT _MF_FAT	GO:0 dmeC GO:0 GO:0 GO:0 CG16 Enric threc IPRO GO:0 GO:0 GO:0	2380me 000502 03050:P 004175 006508 070011 6916, CG hment \$ onine pro 01353:P 004298 070003 005839	oteasome  rendopeptidase activity  rproteolysis  rpeptidase activity, actin 5378, CG18174, CG4157  icore: 2,37  otease  roteasome, subunit alph  rthreonine-type endoper	a/beta otidase activity			7 7 7 7 8 7	19 1 19 1 19 1 19 1 22 2 19 4 8 1 8 3 8 5 8 5	,61E-09 ,03E-06 ,37E-06 ,21E-02 ,39E-02 ,36E-02 ,72E-03 ,46E-03 ,05E-03 ,05E-03	
	KEGG_PA GOTERM_ GOTERM_ GENES Annotation SP_PIR_K INTERPRO GOTERM_ GOTERM_ GOTERM_ GOTERM_ GENES	CC_FAT THWAY _MF_FAT _BP_FAT _MF_FAT on Cluster 2 EYWORDS ) _MF_FAT _MF_FAT	GO:0 dmeC GO:0 GO:0 GO:0 CG16 Enric threc IPR00 GO:0 GO:0 CG49	easome 000502 03050:P 004175 006508 070011 5916, Co hment \$ onine pro 01353:P 004298 070003 005839	oteasome  control of the control of	a/beta otidase activity			7 7 7 7 8 7	19 1 19 1 19 1 19 1 22 2 19 4 8 1 8 3 8 5 8 5	,61E-09 ,03E-06 ,37E-06 ,21E-02 ,39E-02 ,36E-02 ,72E-03 ,46E-03 ,05E-03 ,05E-03	
	KEGG_PA GOTERM_ GOTERM_ GENES Annotation SP_PIR_K INTERPRO GOTERM_ GOTERM_ GOTERM_ GOTERM_ GENES	CC_FAT THWAY _MF_FAT _BP_FAT _MF_FAT on Cluster 2 EYWORDS ) _MF_FAT _MF_FAT _CC_FAT on Cluster 7	GO:0 dmeC GO:0 GO:0 GO:0 CG16 Enric three IPR00 GO:0 GO:0 CG49 Enric	easome 000502 03050:P 004175 006508 070011 5916, Co hment \$ onine pro 01353:P 004298 070003 005839 904, CG4	oteasome "endopeptidase activity "proteolysis "peptidase activity, actin 5378, CG18174, CG4157 dore: 2,37 otease roteasome, subunit alph "threonine-type endoper "threonine-type peptidas "proteasome core compl 097, CG12323	, CG4904, CG17 na/beta utidase activity se activity ex			7 7 7 7 8 7	19 1 19 1 19 1 19 1 22 2 19 4 8 1 8 3 8 5 8 5	,61E-09 ,03E-06 ,37E-06 ,21E-02 ,39E-02 ,36E-02 ,72E-03 ,46E-03 ,05E-03 ,05E-03	
	KEGG_PA' GOTERM_ GOTERM_ GOTERM_ Genes Annotatic SP_PIR_K INTERPRC GOTERM_ GOTERM_ GOTERM_ GOTERM_ GENES Annotati	CC_FAT THWAY MF_FAT BP_FAT MF_FAT On Cluster 2 EYWORDS ) MF_FAT MF_FAT CC_FAT on Cluster 7	GO:0 dmeC GO:0 GO:0 GO:0 CG16 Enric three IPRO0 GO:0 GO:0 GO:0 GO:0 GO:0 GO:0 GO:0 GO	easome 000502 03050:P 004175 006508 070011 6916, CG hment \$ 001353:P 004298 070003 005839 904, CG4 hment \$	oteasome "endopeptidase activity "proteolysis "peptidase activity, actin 5378, CG18174, CG4157 doce: 2,37 otease roteasome, subunit alph "threonine-type endoper "threonine-type peptidas "proteasome core compl 1097, CG12323 Gore: 1,92	, CG4904, CG17  aa/beta bitidase activity se activity ex  on organization			7 7 7 7 8 7 3 3 3 3 3 3	19 1 19 1 19 1 19 1 22 2 19 4 8 1 8 3 8 5 8 9	,61E-09 ,03E-06 ,37E-06 ,21E-02 ,39E-02 ,36E-02 ,72E-03 ,46E-03 ,05E-03 ,76E-03	
	KEGG_PA GOTERM GOTERM GOTERM GENES Annotati SP_PIR_K INTERPRC GOTERM	CC_FAT THWAY MF_FAT BP_FAT MF_FAT On Cluster 2 EYWORDS ) MF_FAT MF_FAT CC_FAT  On Cluster 7 BP_FAT BP_FAT BP_FAT	GO:0 dmeC GO:0 GO:0 GO:0 GO:0 Enric IPRO0 GO:0 GO:0 GO:0 GO:0 GO:0 GO:0 GO:0 GO	easome 000502 03050:P 004175 006508 070011 9916, CG hment S 011853:P 004298 070003 005839 004, CG chment S	roteasome rendopeptidase activity rproteolysis repetidase activity, actin 5378, CG18174, CG4157 core: 2,37 otease roteasome, subunit alph rthreonine-type endopep rthreonine-type peptidas rproteasome core compl 097, CG12323 core: 1,92 rmicrotubule cytoskelet rmitotic spindle organiz rspindle organization	a/beta altidase activity ex on organization ation			7 7 7 7 8 7 8 7 3 3 3 3 3 3 3 5	19 1 19 1 19 1 19 1 22 2 19 4 8 1 8 3 8 5 8 9	,61E-09 ,03E-06 ,37E-06 ,21E-02 ,39E-02 ,36E-02 ,72E-03 ,46E-03 ,05E-03 ,76E-03 ,76E-03 ,28E-03 ,21E-02	
	KEGG_PA GOTERM GOTERM GOTERM GENES ANNOTATION GOTERM	CC_FAT THWAY MF_FAT BP_FAT MF_FAT On Cluster 2 EYWORDS ) MF_FAT MF_FAT CC_FAT  On Cluster 7 BP_FAT BP_FAT BP_FAT	GO:0 dme0 GO:0 GO:0 GO:0 GO:0 Enric threc IPR00 GO:0 GO:0 GO:0 GO:0 GO:0 GO:0 GO:0 GO	easome 000502 03050:P 004175 006508 070011 9916, CG hment S onine pro 01353:P 004298 070003 005839 904, CG 404000226 0000226 0007051	oteasome  rendopeptidase activity renoteolysis repetidase activity, actin 5378, CG18174, CG4157 core: 2,37  otease roteasome, subunit alph rethreonine-type endopep rethreonine-type peptidas reproteasome core compl 097, CG12323 core: 1,92  remicrotubule cytoskelet remitotic spindle organiz	a/beta utidase activity es on organization on	82,CG4097, CG12		7 7 7 7 8 7 3 3 3 3 3 3	19 1 19 1 19 1 19 1 22 2 19 4 8 1 8 3 8 5 8 9	,61E-09 ,03E-06 ,37E-06 ,21E-02 ,39E-02 ,36E-02 ,72E-03 ,46E-03 ,05E-03 ,76E-03 ,76E-03	

**Fig. 3** GO term clustering reveals the proteasome as efficient target. **a** All 11 top RNAi target genes have *Drosophila* orthologs and overall, their protein sequence identity with other species is high. Only Cact, Inr-a and Gw show a low degree of sequence conservation. Dm *Drosophila melanogaster*, Am *Apis mellifera*, Aa *Aedes aegypti*, Ap *Acyrthosiphon pisum*. **b** GO term clustering of the top 40 RNAi target genes revealed ten clusters with enriched biological processes. Searching for genes that share the respective GO term combinations identified additional RNAi target genes. For cluster 1, seven novel genes with the respective combination were found and six of them (86 %) turned out to be highly lethal at day 11 (d11) after injection. Cluster 1, 2 and 7 showed the highest predictive power. **c** The GO terms of the most predictive clusters are shown. *Count* is the number of genes annotated with a given term.

GOTERM BP FAT are annotations with respect to biological processes (see Additional file 2: Figure S4 for all clusters)

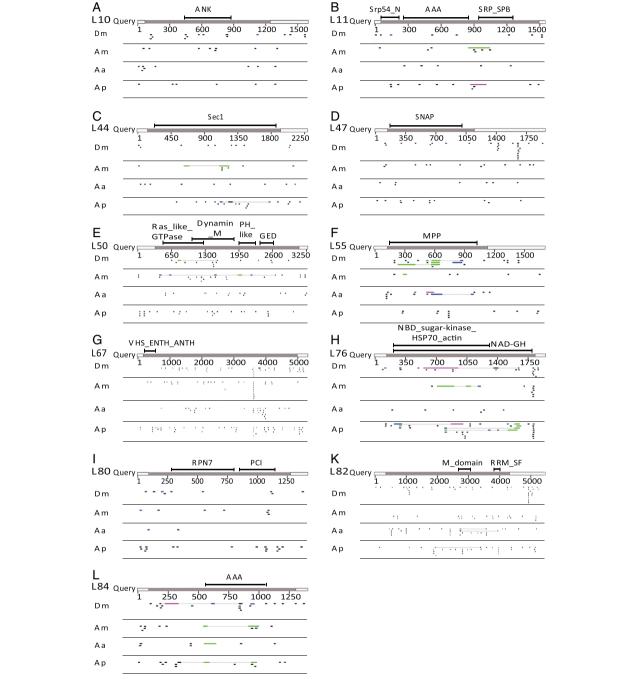
habitats, it will be difficult if not impossible to identify target sequences without potential off-target site in any other species.

Note that most of the identified potential 17 bp off-target-regions will not lead to any RNAi response and that at least half of the potentially targeted genes will not lead to lethality [24]. Further, only individuals that actively eat the protected plants will suffer. Hence, the species specificity of the RNAi based technique remains unchallenged when compared to other methods that usually target all or at least many species. Nevertheless, efforts to increase safety should

focus on selected species (like bees) that need to be protected in the respective given ecological setting.

## GO term cluster identifies the proteasome as prime RNAi target gene

We tried to identify common properties of our identified RNAi targets, because this information might help identifying novel RNAi targets in species less amenable to large scale screens. We first analyzed the adult body expression levels in *Tribolium* and compared them to their *Drosophila* orthologs [27]. A striking pattern was not found apart from



**Fig. 4** Location of potential off target regions is not restricted to stretches of conserved protein sequence. **a-I** The non-coding and rapidly evolving UTRs are indicated as *open bars*, the coding sequence as *grey bar*. Conserved protein domains are indicated above the coding sequence by *black bars* and the protein domain name. Sites with ≥17 nt identity were identified. The hits were plotted at the respective position of the query for each species (rows below the query). We did not find strong if any correlation between the degree of sequence conservation (UTR < coding sequence < conserved domains) and the location of potential off target sites. This indicates that it will be difficult if not impossible to exclude off target effects when misexpressing dsRNAs in plants. Multiple isoforms of different genes were not figured. Dm *Drosophila melanogaster*, Am *Apis mellifera*, Aa *Aedes aegypti*, Ap *Acyrthosiphon pisum* 

generally high expression of the *Drosophila* orthologs, which was specifically true for the central nervous system (Additional file 2: Figure S3 M).

Next, we searched for GO term clusters of the top 11 and top 40 RNAi targets [28, 29] using Drosophila melanogaster GO term annotations as the background. The clustering of the 40 RNAi target genes resulted in 10 clusters (Additional file 2: Figure S4). In order to test, in how far these GO terms are predictive, we identified 1328 Drosophila genes sharing the same GO term combinations of at least one of these clusters. For those, we determined the respective Tribolium orthologs and found that 502 of them had by chance been included in the iBeetle screen. Almost all novel genes identified by GO term combinations of cluster 1 and 2 (GO terms related to proteasome function) showed a strong lethality in the iBeetle screen (≥70 % after pupal or larval injection; Fig. 3b, c). Cluster 7 (GO terms related to cytoskeleton organization) represented the thirdbest cluster with 37 of 66 novel genes showing lethality in the screen (Fig. 3b, c). The clustering of the top 11 RNAi target genes did not result in clusters, which were able to predict novel RNAi target genes, which could be due to the low number of input genes, which makes statistical analysis challenging (not shown). Taken together, this analysis reveals GO term combinations that are predictive for potential RNAi target genes and reveal the proteasome as prime target for RNAi based insecticides.

### Conclusion

We extend the number of potential RNAi target genes by two means: First, we used the red flour beetle T. castaneum as screening platform in order to identify RNAi target genes by the hypothesis independent large scale RNAi screen iBeetle [24]. Indeed, novel RNAi target genes were revealed that efficiently induced death in the pest species Tribolium castaneum and that performed better compared to previously used genes. Hence, the RNAi target genes identified in this study are excellent candidates in future efforts in other pest species. Secondly, we identify GO term clusters that are predictive for genes with lethal phenotype upon RNAi. The search for the most lethal genes in e.g. gut transcriptomes of pest species will be accelerated by selecting genes that show these GO term combinations. Third, the proteasome appears to be a prime target for pest control. At this stage we cannot exclude that oral intake of dsRNA will have different dynamics than dsRNA injection. Hence, the tests in the pest species will have to include oral feeding assays.

It had remained unclear whether plant protection efficiency would increase by expressing a dsRNA fragment targeting two target genes at the same time. Despite extensive testing we did find no a synergism between our most effective RNAi target genes.

Further, our analysis indicates that the number of off target sites does not strongly correlate with protein conservation and that it will be difficult if not impossible to design dsRNA fragments that do not have off target effects in any non-target species. One conclusion is that the search for RNAi target genes does not need to be restricted to species or taxon specific genes. The second conclusion is that the design of the target sequences needs to focus on avoiding off target effects on selected species that should be protected (like for instance bees) or that are most common in the respective environment.

### **Methods**

#### Strains

San Bernadino strain (SB) was used and reared on wholegrain flour supplemented with 5 % yeast powder at 32 °C for all experiments.

#### **RNAi**

All dsRNAs were ordered from Eupheria Biotech GmbH (www.eupheria.com), titrated from 0,5 ng/µl to 1 µg/µl and approximately 0.15 µl were injected into 10 larvae or adult beetles, respectively, using established techniques [30]. For the best 40 target genes non-overlapping fragments (1 µg/µl) were injected in additional RNAi experiments as a control for off-target effects. RNAi targeting the pigmentation gene Tc-ebony was used as negative control in all experiments at the respective concentrations. The change in body color confirmed successful RNAi while the survival rate was documented as negative control of the experiment. In the iBeetle screen 251 buffer injections were added as buffer control with 97 % not displaying any phenotype [24]. The survival rate was scored every second day post treatment by counting those animals that still showed movements. This number was divided by the number of injected animals in order to calculate the percentage of survival. All primers and sequences are documented at iBeetle Base [23].

# GO term clustering and identification of novel potential RNAi target genes

DAVID 6.7 (The Database for Annotation, Visualization and Integrated Discovery) [28, 29] was used to analyse the 11 and 40 RNAi most efficient target genes. The Drosophila genome was set as background and p-values (p-value  $\leq$  0.05) represented a modified Fisher's exact t-test. The enriched GO terms were clustered into classes by the functional annotation clustering tool, which uses a grouping algorithm based on the hypothesis that similar genes should share similar annotations. The option settings were: classification stringency "high" and enrichment thresholds "EASE 0.05". The enrichment score of each group is the geometric mean (in-log) of the p-values in an annotation cluster. Thus, the uppermost group shows the highest

biological significance. GO Fat category was used for this analysis, which is a subset comprising more specific terms. The clustering of the eleven RNAi target genes was done with ten genes in total, because L67 was not associated with any GO term. For the top 40 clustering, 37 genes were used, due to three genes without associated GO terms. Note that the last DAVID GO database update was in 2009, some GO terms have changed in the last years and could be wrongly annotated in the clusters. In order to find further potential RNAi target genes, the GO terms of each cluster were used in the Flybase Query Builder [27]. With the obtained Flybase gene IDs we searched for Tribolium orthologs in the iBeetle Base [23]. Genes that caused a mortality of ≥70 % in the screen on day 11 after pupal or larval injection were assumed to be novel potential RNAi target genes.

### Off target analysis

The nucleotide sequence of the RNAi target genes (Query) was used to identify potential off target sites in transcript sequences of other species by Blast analysis. To this end, the length of the exact match was defined as ≥15 nt by the word size function at NCBI Blast and the match/mismatch score was defined as 1/–4. Exact matches smaller than 17 nt were excluded manually.

### Ethics statement

All experiments have been performed with invertebrate animals according to the respective institutional, national, and international guidelines (Directive 2010/63/EU). Approval by an ethics committee was not required.

### Additional files

**Additional file 1: Table S1.** lists the top 100 RNAi target genes with their *Tribolium* iBeetle-numbers, *Tc*-genes and *Drosophila* orthologs. **Table S2.** lists the best 40 RNAi target genes with their *Tribolium* iBeetle-numbers, *Tc*-genes and *Drosophila* orthologs. **Table S3.** lists the number of surviving animals after injection with non-overlapping dsRNA fragments. **Table S4.** lists the number of surviving larvae after single and double RNAi treatments (total dsRNA concentration 0,5 ng/ $\mu$ l). **Table S5.** shows the sequences of the *Tribolium* orthologs of the RNAi target genes published by Baum et al. [15]. (XLSX 37 kb)

**Additional file 2: Figure S1.** shows the results for the most efficient 40 RNAi target genes. In **Figure S2.** the results of dsRNA injection into adult beetles are displayed. **Figure S3.** shows the phylogenetic trees of the novel RNAi target genes. **Figure S4.** shows the GO term clusters of the top 40 RNAi target genes. (PDF 3696 kb)

### Competing interests

AV and GB cooperate to use the presented data to establish transgenic oilseed rape plants protected against the pollen beetle. This publication has no influence on any kind of remuneration or funding of the project.

### Authors' contributions

JU did all *Tribolium* experiments and sequence analyses, contributed to the GO term analysis and drafted the manuscript. CSE, UM, JS, DS, NS, NT, VAD, DG, TR identified the lethality of the respective genes in the iBeetle screen. JD and LG contributed to the GO term analysis. LG and MS did the

expression analysis. GL designed and supervised RNAi experiments in other species. AV drafted parts of the manuscript. GB conceived of the study, and participated in its design and coordination and drafted the manuscript. All authors read and approved the final manuscript.

### Acknowledgments

We acknowledge support by the German Research Foundation (DFG), which funded the *iBeetle* project (FOR1234) but was not involved in design or analysis of the data or its publication. Further, we acknowledge the Open Access Publication Funds of the Göttingen University. We thank Claudia Hinners for technical assistance, Maike Tech for extracting the gene lists from the database.

### **Author details**

<sup>1</sup>Johann-Friedrich-Blumenbach-Institut, GZMB, Georg-August-Universität Göttingen, Justus-von-Liebig-Weg 11, 37077 Göttingen, Germany. <sup>2</sup>Department Biologie, Entwicklungsbiologie, Friedrich-Alexander-Universität Erlangen-Nürnberg, Staudtstraße 5, 91058 Erlangen, Germany. <sup>3</sup>Institut für Entwicklungsbiologie, Universität zu Köln, Zülpicher Straße 47b, 50674 Cologne, Germany. <sup>4</sup>Institut für Mathematik und Informatik, Ernst Moritz Arndt Universität Greifswald, Walther-Rathenau-Straße 47, 17487 Greifswald, Germany. <sup>5</sup>Institut für Biologie, Neurobiologie, Freie Universität Berlin, Königin-Luise-Str. 28/30, 14195 Berlin, Germany. <sup>6</sup>Institut für Phytopathologie und Angewandte Zoologie, Justus-Liebig-Universität Gießen, Heinrich-Buff-Ring 26-32, D-35392 Gießen, Germany.

Received: 23 February 2015 Accepted: 26 August 2015 Published online: 03 September 2015

#### References

- 1. Moffat AS. Finding new ways to fight plant diseases. Science. 2001;292:2270-3.
- Ffrench-Constant RH. The molecular genetics of insecticide resistance. Genetics. 2013;194:807–15.
- Ffrench-Constant RH. Insecticide resistance comes of age. Genome Biol. 2014:15:106
- 4. Tabashnik BE, Brévault T, Carrière Y. Insect resistance to Bt crops: lessons from the first billion acres. Nat Biotechnol. 2013;31:510–21.
- Price DRG, Gatehouse JA. RNAi-mediated crop protection against insects. Trends Biotechnol. 2008;26:393–400.
- Narva KE, Siegfried BD, Storer NP. Transgenic approaches to western corn rootworm control. Adv Biochem Eng Biotechnol. 2013;136:135–62.
- Scott JG, Michel K, Bartholomay LC, Siegfried BD, Hunter WB, Smagghe G, et al. Towards the elements of successful insect RNAi. J Insect Physiol. 2013;59:1212–21.
- Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature. 1998;391:806–11.
- Meister G, Tuschl T. Mechanisms of gene silencing by double-stranded RNA. Nature. 2004;431:343–9.
- Brown SJ, Mahaffey JP, Lorenzen MD, Denell RE, Mahaffey JW. Using RNAi to investigate orthologous homeotic gene function during development of distantly related insects. Evol Dev. 1999;1:11–5.
- Bucher G, Scholten J, Klingler M. Parental RNAi in Tribolium (Coleoptera). Curr Biol. 2002;12:R85–6.
- Tomoyasu Y, Denell RE. Larval RNAi in Tribolium (Coleoptera) for analyzing adult development. Dev Genes Evol. 2004;214:575–8.
- Tomoyasu Y, Miller SC, Tomita S, Schoppmeier M, Grossmann D, Bucher G. Exploring systemic RNA interference in insects: a genome-wide survey for RNAi genes in Tribolium. Genome Biol. 2008;9:R10.
- Whangbo JS, Hunter CP. Environmental RNA interference. Trends Genet TIG. 2008;24:297–305.
- Baum JA, Bogaert T, Clinton W, Heck GR, Feldmann P, Ilagan O, et al. Control of coleopteran insect pests through RNA interference. Nat Biotechnol. 2007;25:1322–6.
- Whyard S, Singh AD, Wong S. Ingested double-stranded RNAs can act as species-specific insecticides. Insect Biochem Mol Biol. 2009;39:824–32.
- Miller SC, Brown SJ, Tomoyasu Y. Larval RNAi in Drosophila? Dev Genes Evol. 2008;218:505–10.
- Berghammer AJ, Klingler M, Wimmer EA. A universal marker for transgenic insects. Nature. 1999;402:370–1.

- Trauner J, Schinko J, Lorenzen MD, Shippy TD, Wimmer EA, Beeman RW, et al. Large-scale insertional mutagenesis of a coleopteran stored grain pest, the red flour beetle Tribolium castaneum, identifies embryonic lethal mutations and enhancer traps. BMC Biol. 2009;7:73.
- Schinko J, Hillebrand K, Bucher G. Heat shock-mediated misexpression of genes in the beetle Tribolium castaneum. Dev Genes Evol. 2012;222(5):287–98.
- Schinko JB, Weber M, Viktorinova I, Kiupakis A, Averof M, Klingler M, et al. Functionality of the GAL4/UAS system in Tribolium requires the use of endogenous core promoters. BMC Dev Biol. 2010;10:53.
- Richards S, Gibbs RA, Weinstock GM, Brown SJ, Denell R, Beeman RW, et al. The genome of the model beetle and pest Tribolium castaneum. Nature. 2008;452:949–55.
- Dönitz J, Schmitt-Engel C, Grossmann D, Gerischer L, Tech M, Schoppmeier M, et al. iBeetle-Base: a database for RNAi phenotypes in the red flour beetle Tribolium castaneum. Nucleic Acids Res. 2015;43:D720–5.
- Schmitt-Engel C, Schultheis D, Schwirz J, Ströhlein N, Troelenberg N, Majumdar U et al. The iBeetle large-scale RNAi screen reveals gene functions for insect development and physiology. Nat Commun. 2015;6:7822. doi:10.1038/ncomms8822.
- Kulkarni MM, Booker M, Silver SJ, Friedman A, Hong P, Perrimon N, et al. Evidence of off-target effects associated with long dsRNAs in Drosophila melanogaster cell-based assays. Nat Methods. 2006;3:833–8.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10.
- Gelbart WM, Crosby MA, Matthews B, Rindone WP, Chillemi J, Twombly SR, et al. FlyBase: a Drosophila database. The FlyBase consortium. Nucleic Acids Res. 1997:25:63–6.
- Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4:44–57.
- Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009;37:1–13.
- Posnien N, Schinko J, Grossmann D, Shippy TD, Konopova B, Bucher G. RNAi in the red flour beetle (Tribolium). Cold Spring Harb Protoc. 2009(8):pdb.prot5256.

## Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit

