## Molecular cloning and transcriptional activation of lysozyme-encoding cDNAs in the mosquito Aedes aegypti

## R. J. Ursic Bedoya\*, A. M. Mitzey†, M. Obraztsova‡ and C. Lowenberger\*

\* Department of Biological Sciences, Simon Fraser University, 8888 University Dr., Burnaby, BC, Canada; † School of Veterinary Medicine, University of Wisconsin-Madison, 2015 Linden Dr., Madison WI, USA; and ‡Biotechnology Center, University of Wisconsin-Madison, 425 Henry Mall, Madison WI, USA

## Abstract

Lysozymes are enzymes characterized by their ability to break down bacterial cell walls. In insects certain lysozymes are only found in the midgut, whereas others are only found in the haemolymph and fat body after immune challenge. We identified two lysozyme-encoding cDNAs from Aedes aegypti. Both deduced protein sequences are basic in nature, contain 148 amino acids including eight highly conserved cysteine residues, and their genomic sequences contain a single intron. Transcriptional profiles indicated that the predominant form is constitutively expressed and up-regulated upon immune challenge and blood feeding in adult mosquitoes. The second form is expressed during early developmental stages, larvae and pupae, and at low levels in adults after immune challenge. Lysozymes in Aedes aegypti play both roles, defined by the spatial and temporal regulation of their expression.

## Keywords: lysozyme, insect immunity.

## Introduction

Insects have exploited almost every niche on Earth due in part to their ability to mount a rapid and potent immune response to limit pathogen development. Insects lack the acquired immune response found in higher organisms but possess a well-developed innate response. Cellular

doi: 10.1111/j.1365-2583.2004.00534.x

mediated defences include phagocytosis, nodulation or encapsulation of bacterial pathogens (Lavine & Strand, 2002; Hillyer *et al.*, 2003), or the melanization of larger organisms that may be killed directly or by the expression of reactive intermediates of nitrogen or oxygen (Nappi & Vass, 2001). Humoral defences are characterized by an extremely rapid induction of immune peptides: transcripts for immune peptides are found within minutes of bacterial insult, and may last for weeks (Lowenberger, 2001).

Studies of mosquito immune responses have looked for novel molecules expressed specifically as a defence mechanism to limit the development of human pathogens. Less well studied is the role of constitutively expressed molecules, which may be used in more than one role.

Lysozymes, present in a large range of organisms, are defined by their enzymatic cleavage of the  $\beta$ 1,4-glycosidic linkage between N-acetylmuramic acid and N-acetylglucosamine found in certain bacterial cell walls. Lysozyme function has been widely described in species that ingest or harbour bacteria throughout their life cycles, such as Musca domestica and Drosophila melanogaster, and are characterized as digestive enzymes (Regel et al., 1998; Terra & Espinoza-Fuentes, 1987). In mosquitoes, lysozymes have been classified as important components of the immune system (Gao & Fallon, 2000). We hypothesize that lysozymes play a dual role: one in a digestive manner to break down ingested bacteria in the gut, and two as a defence response against pathogens that enter the haemocoel. We describe here a novel lysozyme cDNA and propose defensive and digestive roles for lysozymes in Ae. aegypti.

## Results

## cDNA clones for Ae. aegypti lysozymes

We obtained the full length of *Ae. aegypti* Lys-A, previously identified from an immune challenged cell line (Gao & Fallon, 2000), and the gDNA of which contains a 74 bp intron. We obtained a novel lysozyme cDNA, designated Lys-E (GenBank accession number AY693973), comprising 636 bp with a 5' untranslated region of fifty-five bases, a start codon ATG, 444 bases that correspond to a protein of 148 amino acids, a stop codon, forty-five bases of 3' untranslated region, the putative polyadenylation signal (AATAAA), 83

Received 7 April 2004; accepted after revision 24 September 2004. Correspondence: Raul J. Ursic Bedoya, Biological Sciences Department, Simon Fraser University, 8888 University Dr., Burnaby, BC V5A1S6, Canada. Tel.: +1 604 291 4391; e-mail: rursicbe@sfu.ca

 $\verb|ccaaagactgattcaaacagtccgtgcgaaaaatcggtgttcaaacaccatcaga||$ 

Ρ v v K G N Α Μ Α ATG GGA AAC GCA AAG TTG CCG ATC GTA CTG GTA GGA CTT TTT GTC CTA CTG GCA TCC ACA Κ F D K С E Α Κ S L Α Κ Α L T. Α 0 G ATC GAA GCA AAA AAG TTC GAC AAA TGT TCG CTG GCC AAG GCT TTG CTG GCT CAA GGA TTC К Α D T. R N W V C T. V 0 N E S S М D S т AGC AAA GCA GAT CTA CGT AAC TGG GTC TGT CTG GTG CAG AAT GAA AGC AGC ATG GAC ACC Κ Η Ν Ν S Ν G S Т D W G L F 0 AAC AAG AAG CAC AAC AAC AGC AAT GGA TCC ACC GAT TGG GGC CTG TTC CAG ATC AAC GAT R W C D Ρ 0 D Κ K Κ Κ Т S Ν E C Κ Κ CGA TAC TGG TGC GAT CCA CAG GAC AAG AAG AAA AAG ACC TCG AAC GAG TGT AAG GTC AAA C F K S D Ι S Κ Α S т C K S E Τ. Δ K TGT TCG G^AA TTC CTG AAA AGT GAC ATC AGC AAG GCT TCC ACA TGT GCG AAA AAA ATC TAC S W Y G W т Ν Κ C Н G F Α Κ G K 77 AAA CGA CAC GGT TTT AGC GCC TGG TAC GGA TGG ACA AAC AAA TGT AAA GGC AAA GTT CTT D L S K С K L CCC GAT TTG AGC AAA TGT AAA CTG TGA aaaatqqqqacaaacaaataatqtttcqaqacaaataatactatqaataaatttacactatacttctqqtatttqaaa 

 $introm\ sequence:\ gtgagttgtagggatggaaggggttaaattgaaataactgttagataatcatgatgttgttcgaatttttag$ 

Figure 1. cDNA and translated amino acid sequences of *Aedes aegypti* lysozyme-E sequence (GenBank accession number AY693973). The ATG start codon and the TGA stop codon are indicated in bold. The underlined sequence indicates the putative signal peptide region, and the putative polyadenylation consensus sequence is italicized. The intron found in the genomic clone occurs in E (103) and is indicated (Λ).

bases and the poly A tail (Fig. 1). The peptide is produced as a preprotein and is activated by a cleavage event. There is an intron of seventy-two bases in the genomic sequence (Fig. 1).

Screening and sequencing of clones isolated from a genomic phage library yielded the two sequences described as Lys-A and -E. Sequencing up to 5 kb upstream of the coding region from both genes (data not shown) revealed the existence of a TATA box around position –72 and a NF- $\kappa$ B-like regulatory motif which are known to regulate immune genes in *Drosophila melanogaster* (Engstrom *et al.*, 1993). At the amino acid level, these two *Ae. aegypti* lysozymes differ by twenty-three residues, representing 83% shared identity, and are significantly different from lysozyme D found in the salivary glands (Valenzuela *et al.*, 2002) (43% shared identity).

# Deduced physico-chemical properties and similarity to other lysozymes

Aedes aegypti Lys-E and -A have twenty-seven and twentyeight strongly basic residues (K and R), respectively, accounting for nearly 18% of the total number of residues; this confers the basic nature of these proteins with an estimated isoelectric point of 9.40 and 9.45. The N terminus of each protein is hydrophobic, suggesting the presence of a signal peptide, with a predicted cleavage site between positions A23 and K24 (EA-KK) (SignalP v.1.1). Cleavage of the signal peptide produces active peptides of 14.3 and 14.4 kDa for Lys-E and -A, respectively, but does not alter their isoelectric points. Protein sequence alignment with other lysozymes from selected organisms (Fig. 2a) indicates they belong to the chicken c-type, with eight highly conserved cysteine residues at positions 29, 50, 84, 97, 101, 115, 135 and 146, the conserved active site glutamic acid (E55) and aspartic acid (D72).

#### Phylogenetic analysis of insect lysozymes

cDNA and protein sequences were used to determine the closest associations among insect lysozymes (Fig. 2b). The data suggest an early divergence of lysozymes from a common ancestor of Lepidoptera and mosquitoes. We find a dichotomy between recruitment of lysozyme for a digestive function (mainly in Cyclorrhapha flies) and as an immune peptide (Lepidoptera). There also is a divergence within the mosquitoes in that the salivary gland lysozyme is quite distinct from the other mosquito lysozymes. It has been suggested that this lysozyme serves to control the microbial population ingested with a sugar meal (Rossignol & Lueders, 1986). Although Lys-E and -A are expressed in the midgut, an exclusive recruitment of these lysozymes to a digestive function has not occurred: they are more closely related to Lepidoptera immune genes than to cyclorrhaphan digestive lysozymes or anopheline salivary gland lysozymes.

## Developmental, tissue specificity and induction of lysozyme transcription in Ae. aegypti

We used real-time quantitative PCR (Q-PCR) to quantify spatial and temporal expression of lysozyme transcripts. Both lysozymes described here are found in all samples but Lys-A transcripts were far more abundant. Transcripts for both lysozymes were found in low levels in larvae and white pupae, and only in these stages did transcript numbers of Lys-E exceed those of Lys-A. In black pupae there was a fifteen-fold increase in Lys-A transcripts but no significant difference in Lys-E transcripts over levels found in white pupae. There was a three-fold greater number of Lys-A over Lys-E transcripts in recently emerged adults. Inoculation of bacteria into the haemocoel induced a three-fold increase



Figure 2. (A) Alignment of *Aedes aegypti* lysozymes with other selected insect lysozyme sequences. Gaps used to maintain optimal alignment are indicated by dashes. Alignments were made using MEGALIGN (DNASTAR) by J. Hein method PAM 250. (B) Phylogenetic analysis of selected insect lysozymes. Topologies were obtained using PAUP v.4.0b10 (Swofford, 2002). Accelerated likelihood explorations were performed using likelihood ratchets (Vos, 2003) in conjunction with PAUP Rat (Sikes & Lewis, 2001). A strict consensus of the trees with the same best likelihood was considered the best phylogenetic tree. The sequences used in the upper and lower panels are: *Ae. aegypti* A (gi|12044523|), *Ae. aegypti* B (gi|42764190|), *Ae. aegypti* C (gi|42763735|), *Ae. aegypti* D (gi|18568288), *Ae. albopictus* (gi|20159661|), *Armigeres subalbatus* (gi|42765599|), *Anopheles darlingi* (gi|2198832|), *Anopheles gambiae* (gi|2133628|), *Drosophila melanogaster* (gi|17975524|, gi|17136660|, gi|17136668|, and gi|17136652|), *Musca domestica* (gi|33504660| and gi|32504668|), *Hyalophora cecropia* (gi|159204|), *Samia cynthia* (gi|12082297|), *H. virescens* (gi|4097238|), *Bombyx mori* (gi|567099|) and *Triatoma infestans* (gi|32454476|).



**Figure 3.** Real-time quantitative PCR (Q-PCR) profile of *Aedes aegypti* lysozymes-A and -E in whole bodies or specific tissues after stimulation. Standard curves were established for Lys-A, Lys-E and actin. Products from each sample were normalized to actin levels in each cDNA. The bar charts (mean + SD) represent differences in lysozyme transcripts over an average of four replicates. Lanes: 1, 4th instar larvae; 2, white pupae; 3, black pupae; 4, adult naïve; 5, adult 24 h post inoculation with bacteria; 6, fat body dissected from naïve adults; 7, fat body dissected 24 h post inoculation with bacteria; 8, midguts dissected from adults 24 h post inoculation with bacteria; 10, midguts from adults 2 days post blood-feeding; 11, midguts from adults 5 days post blood-feeding.

in Lys-A transcription but a significant decrease in Lys-E expression (Fig. 3). In dissected tissues we find constitutive fat body expression for both lysozymes. After bacterial inoculation Lys-A transcripts increased nine-fold in the fat body whereas Lys-E expression did not change significantly. In the midgut, bacterial inoculation induced a nine-fold increase in Lys-A transcription, whereas there was no significant change in Lys-E transcripts (Fig. 3). Blood-feeding induced significant expression of Lys-A transcripts over

naïve insects or expression levels of Lys-E, but returned to background levels within 5 days.

Lysozyme transcripts were found in all tissues; Lys-E transcripts were found at low levels, with the highest relative amounts found in larvae and black pupae. Transcripts were found in adults or specific tissues in response to either bacterial challenge or blood-feeding (Fig. 3b), but not to the same level as for Lys-A. The expression of Lys-E seems to be similar to that of Lys-X in *D. melanogaster*, which is expressed in late larvae and early pupal midguts (Daffre *et al.*, 1994).

#### Discussion

In insects, lysozymes have been identified from larval and adult digestive tracts and from salivary glands, suggesting their role in digestion, whereas other reports have indicated a major role in immune responses to pathogens. The data presented here suggest a dual role of two closely related but differentially expressed lysozyme genes in Ae. aegypti. A primary digestive role is deduced for Lys-E (GenBank accession number AY693973) during early developmental stages where larvae primary feed on bacteria, a phenomenon also observed in other insect larvae. In cyclorrhaphan flies such as D. melanogaster and Musca domestica that routinely feed on decomposing material containing bacteria, midgut lysozymes may serve as digestive enzymes (Lemos & Terra, 1991; Regel et al., 1998), and these enzymes may have been modified to operate optimally at the acidic pH of the midgut environment. In addition D. melanogaster lysozymes are not expressed significantly in the fat body, and are not induced in parallel with previously identified immune peptides (Daffre et al., 1994). However, five D. melanogaster lysozymes were shown to be up-regulated after oral infection by a protozoan parasite (Roxstrom-Lindquist et al., 2004). In many nematocerans, lysozymes are found in the haemolymph but not in the gut (Lemos & Terra, 1991), and thus the recruitment of lysozymes as digestive enzymes, and their adaptation to an acidic midgut (as opposed to the mosquito midgut of pH 9.5-10.1), may have occurred after the divergence of the Cyclorrhapha from the Nematocera. During the metamorphosis of mosquitoes, many larval tissues are histolysed and replaced with adult tissues and structures. Within the pupa there must be provisions to prevent the spread of bacteria from the gut throughout the body. The expression of Lys-E in the larval and pupal stages, as well as defensin expression in pupae, may prevent bacteria-induced septicaemia within these insects. In D. melanogaster Lys-X is only expressed in the midgut of early pupae and may play a similar protective role (Daffre et al., 1994).

Lys-A appears to be the principal lysozyme in adult *Ae. aegypti*. It is constitutively expressed, and is up-regulated in the fat body and in gut tissues upon immune activation

or after a blood meal, suggesting a multifunctional role for this gene. However, the up-regulation of Lys-A after a blood meal in Ae. aegypti contrasts with findings in Anopheles gambiae, in which lysozyme expression increased with sugar feeding but not with blood feeding (Kang et al., 1996). Induction of Lys-A in midgut tissues is also observed after inoculation of bacteria into the haemocoel, suggesting that this enzyme can be exported bidirectionally to the gut and also into the haemolymph, or that signalling cascades communicate between immune-competent tissues allowing for expression of peptides in one tissue (midgut) after stimulation of another (fat body) as has been described for other immune peptides (Lowenberger, 2001). The expression profile of mosquito lysozymes demonstrated here is similar to that of other recognized immune peptides, and the simultaneous expression of several peptides may allow them to act synergistically to eliminate bacteria from the haemocoel: Escherichia coli becomes more susceptible to lysozymes in the presence of attacins and the concentration of cecropin required to kill E. coli drops by an order of magnitude in the presence of lysozyme (Chalk et al., 1994). The synchronized expression of several effector molecules allows the insects to clear potentially lethal microbial infections.

The increase of transcripts in fat body and guts after bacterial inoculation supports our hypothesis that Lys A plays a role mainly as a defence molecule. Surprisingly, we also observe an up-regulation after blood-feeding in adult guts, which may indicate that this gene shares regulatory elements with other genes involved in digestion and could be turned on to control the bacterial population ingested when sugar feeding.

Lysozyme expression is controlled by a complex spatial and temporal gene expression pattern, which defines its functional use in the organism. Our initial data revealed  $\kappa$ B-like motifs in the genomic upstream region of each gene, suggesting a strong similarity with mammalian regulatory mechanisms and a conservation of such mechanisms among organisms. With the sequencing of the *Ae. aegypti* genome underway, and the increasing number of EST libraries being constructed (Bartholomay *et al.*, 2004), it is very likely that additional lysozyme genes in the mosquito will be identified.

#### **Experimental procedures**

#### Mosquito rearing

Aedes aegypti were reared as described previously (Lowenberger et al., 1999). Immature stages (larvae, white pupae and black pupae), and naïve adults, were collected from rearing pans and immediately frozen at -80 °C for subsequent nucleic acid isolation.

#### Immune activation

Adult *Ae. aegypti* were anaesthetized with  $CO_2$ , maintained inactivated on wet ice and immune activated as described previously (Lowenberger *et al.*, 1999).

#### RNA extraction and reverse transcription

Total RNA was extracted from selected tissues or whole bodies of adults or immature stages, and cDNA synthesis was performed as described by Lowenberger *et al.* (1999).

#### cDNA amplification and molecular cloning

Degenerate forward primers designed against consensus sequences for several insect lysozymes were used in a standard PCR reaction of 95 °C (4 min), and thirty-five cycles of 95 °C (10 s), 50 °C (10 s) and 72 °C (30 s) on an Indy Air thermocycler (Idaho Technology). PCR products were size fractionated on a 1.2% LMP agarose gel, and bands were excised from the gel. Ligations, transformation, plasmid preparations and DNA sequencing were performed as described previously (Lowenberger *et al.*, 1999). Because several clones shared significant identity with published lysozyme sequences we designed specific primers that we used in 5' and 3' RACE using the Marathon cDNA synthesis kit (Clontech) to obtain full-length cDNA sequences that were cloned, transformed and sequenced as described above.

#### Sequence and identity analysis

Identity analysis of the deduced amino acid sequence of all sequences was done using BLASTp (http://www.ncbi.nlm.nih.gov/Blast/). Prediction of the signal peptide was performed using SignalP v.1.1 (http://www.cbs.dtu.dk/services/SignalP/output.html).

#### Transcriptional analysis

We used the cDNAs constructed from each developmental stage, tissue or adult treatment group (naïve, bacteria inoculated and blood-fed) in a Q-PCR analysis using a Rotor-Gene 3000 (Corbett Research) to quantify the amount of message for each lysozyme sequence. Initial standard curves were constructed with known concentrations of our purified lysozyme cDNAs. We used actin to standardize our cDNA samples. Standard 25 µl PCR reactions were used but included SYBR-Green (Molecular Probes). Primers used were Lys-E-forward: 5'-ATGGGAAACGCAAAGTTGCC-3' and Lys-A-forward: 5'-TCAGAATGAGAAACTTGAATTTAT-3' and lysozyme Q-reverse: 5'-GTGTCGCTTGTAGATCTTTTC-3'. The PCR protocols used to quantify message were: Lys-A: 95 °C 3 min, and then thirty-five cycles of 95 °C for 15 s, 60 °C for 15 s, 72 °C for 25 s; and Lys-E: 95 °C for 3 min and thirty-five cycles of 95  $^{\circ}\text{C}$  for 15 s, 62  $^{\circ}\text{C}$  for 15 s and 72  $^{\circ}\text{C}$  for 25 s. Quantification, melt curve analysis and sample comparison were performed with the Rotor-Gene v.5 software (Corbett Research). Quantification compared levels of each lysozyme transcript per unit of actin in the same samples. A three-fold difference in transcript numbers was considered significantly different.

#### Phylogenetic analysis

Sequences were aligned using the Jotun Hein method within the Megalign program from DNASTAR with PAM 250 residue weight table. Topologies were obtained using PAUP v.4.0b10 (Swofford, 2002). Accelerated likelihood explorations were performed using likelihood ratchets (Vos, 2003) in conjunction with PAUP Rat (Sikes & Lewis, 2001) A strict consensus of the trees with the same best likelihood was considered the best phylogenetic tree.

## Acknowledgements

We thank L. Christensen and J. Chiles for insect rearing, R. Lewis for phylogenetic analysis and D. Cooper for helpful discussion. This research was funded in part by NIH Grant Al49906 to C.L.

## References

- Bartholomay, L.C., Cho, W.-L., Rocheleau, T.A., Boyle, J.P., Beck, E.T., Fuchs, J.F., Liss, P., Rusch, M., Butler, K.M., Wu, R.C.-C., Lin, S.-P., Kuo, H.-Y., Tsao, I.-Y., Huang, C.-Y., Liu, T.-T., Hsiao, K.-J., Tsai, S.-F., Yang, U.-C., Nappi, A.J., Perna, N.T., Chen, C.-C. and Christensen, B.M. (2004) Description of the transcriptomes of immune response-activated hemocytes from the mosquito vectors *Aedes aegypti* and *Armigeres subalbatus*. *Infect Immun* 72: 4114–4126.
- Chalk, R., Townson, H., Natori, S., Desmond, H. and Ham, P.J. (1994) Purification of an insect defensin from the mosquito, *Aedes aegypti. Insect Biochem Mol Biol* **24**: 403–410.
- Daffre, S., Kylsten, P., Samakovlis, C. and Hultmark, D. (1994) The lysozyme locus in *Drosophila melanogaster*: an expanded gene family adapted for expression in the digestive tract. *Mol Gen Genet* 242: 152–162.
- Engstrom, Y., Kadalayil, L., Sun, S.C., Samakovlis, C., Hultmark, D. and Faye, I. (1993) Kappa B-like motifs regulate the induction of immune genes in *Drosophila*. J Mol Biol 232: 327– 333.
- Gao, Y. and Fallon, A.M. (2000) Immune activation upregulates lysozyme gene expression in *Aedes aegypti* mosquito cell culture. *Insect Mol Biol* 9: 553–558.
- Hillyer, J.F., Schmidt, S.L. and Christensen, B.M. (2003) Rapid phagocytosis and melanization of bacteria and Plasmodium sporozoites by hemocytes of the mosquito *Aedes aegypti*. *J Parasitol* 89: 62–69.
- Kang, D., Romans, P. and Lee, J.Y. (1996) Analysis of a lysozyme gene from the malaria vector mosquito, *Anopheles gambiae*. *Gene* **174**: 239–244.

- Lavine, M.D. and Strand, M.R. (2002) Insect hemocytes and their role in immunity. *Insect Biochem Mol Biol* **32**: 1295– 1309.
- Lemos, F.J. and Terra, W.R. (1991) Digestion of bacteria and the role of midgut lysozyme in some insect larvae. *Comp Biochem Physiol B* **100**: 265–268.
- Lowenberger, C. (2001) Innate immune response of Aedes aegypti. Insect Biochem Mol Biol **31**:219–229.
- Lowenberger, C., Charlet, M., Vizioli, J., Kamal, S., Richman, A., Christensen, B.M. and Bulet, P. (1999) Antimicrobial activity spectrum, cDNA cloning, and mRNA expression of a newly isolated member of the cecropin family from the mosquito vector *Aedes aegypti. J Biol Chem* **274**: 20092–20097.
- Nappi, A.J. and Vass, E. (2001) Cytotoxic reactions associated with insect immunity. *Adv Exp Med Biol* **484**: 329–348.
- Regel, R., Matioli, S.R. and Terra, W.R. (1998) Molecular adaptation of *Drosophila melanogaster* lysozymes to a digestive function. *Insect Biochem Mol Biol* **28**: 309–319.
- Rossignol, P.A. and Lueders, A.M. (1986) Bacteriolytic factor in the salivary glands of *Aedes aegypti. Comp Biochem Physiol B* 83: 819–822.
- Roxstrom-Lindquist, K., Terenius, O. and Faye, I. (2004) Parasitespecific immune response in adult *Drosophila melanogaster*: a genomic study. *EMBO Rep* 5: 207–212.
- Sikes, D.S. and Lewis, P.O. (2001) *PAUPRat: a Tool to Implement Parsimony Ratchet Searches Using PAUP\*, Beta Software,* Version 1. Sinauer Associates, Sunderland, MA.
- Swofford, D.L. (2002) PAUP\* Phylogenetic Analysis Using Parsimony, Version 4.0b10. Sinauer Associates, Sunderland, MA.
- Terra, W.R. and Espinoza-Fuentes, F.P. (1987) Physiological adaptations for digesting bacteria. Water fluxes and distribution of digestive enzymes in *Musca domestica* larval midgut. *Insect Biochem* **17**: 809–817.
- Valenzuela, J.G., Pham, V.M., Garfield, M.K., Francischetti, I.M. and Ribeiro, J.M. (2002) Toward a description of the sialome of the adult female mosquito *Aedes aegypti. Insect Biochem Mol Biol* **32**: 1101–1122.
- Vos, R.A. (2003) Accelerated likelihood surface exploration: the likelihood ratchet. Syst Biol 52: 368–373.