

Molecular Guidelines

The *Journal of Chemical Ecology* has been publishing an increasing number of contributions that report experiments that use the methods of molecular biology. In 2004, a special issue of the Journal [*JCE* 30(12)] focused on molecular chemical ecology and documented the increasing trend of published submissions with molecular content (Seybold 2004). The guidelines presented here were developed to provide assistance to authors, with the goal of ensuring standardized and complete reporting of molecular information in the Journal. The fields of molecular biology and bioinformatics change rapidly as new technologies and protocols are continually developed and introduced. Thus, the Journal intends to revise and re-issue these guidelines periodically as significant changes in this area of science occur. Molecular biological content includes reports on specific nucleic acids and proteins involved in chemical ecological phenomena. The guidelines are separated into 3 parts: I. Types of Papers Published; II. Appropriate Nomenclature and Abbreviations; and III. Presentation of Molecular Biological Content.

I. Types of Papers Published

- A) Isolation, identification, and/or expression of a gene or collections of genes coding for pheromone or odorant binding proteins, olfactory receptors, enzymes involved in pheromone biosynthesis, or genes involved in plant defense (e.g., Xiu and Dong 2007; Zhang et al. 2009; Shivaji et al. 2010).
- B) Characterization of animal molecular, biochemical, or physiological responses to plant defense compounds (e.g., molecular responses to tannins, etc., Da Costa et al. 2008; Wen et al. 2009).
- C) Characterization of plant molecular, biochemical, or physiological responses to animal herbivory (Han et al. 2009; Sethi et al. 2009).
- D) Characterization of peptides and other large biomolecules that serve as olfactory signals (e.g., Robertson et al. 2007).
- E) RNA interference experiments with animals and plants that involve the interruption of chemical ecological phenomena (e.g., Pelletier et al. 2010).
- F) Characterization of all nucleic acid, protein, or natural product responses of organisms to environmental stimuli (genomic, proteomic, or metabolomic studies).

II. Nomenclature and Abbreviations

The following terms and abbreviations should be used in molecular papers submitted to the Journal. Note that the symbols for names of genes and transcripts should be written with italicized font (e.g., *HMG-R* and *HMG-R* mRNA), and should be consistent with the international standards accepted for the organism(s) of study. Symbols for names of proteins should be written in normal font (e.g., 3-hydroxy-3-methylglutaryl-CoA reductase, HMG-R).

A. Abbreviations for Biomolecules and Related Terms

- 1) **Amino Acids:** Ala (A) alanine, Arg (R) arginine, Asn (N) asparagine, Asp (D) aspartic acid or aspartate, Cys (C) cysteine, Gln (Q) glutamine, Glu (E) glutamic acid or glutamate, Gly (G) glycine, His (H) histidine, Hyp (O) hydroxylysine, Hyp (O) hydroxyproline, Ile (I) isoleucine, Leu (L) leucine, Lys (K) lysine, Met (M) methionine, Orn (O) ornithine, Phe (F) phenylalanine, Pro (P) proline, Ser (S) serine, Thr (T) threonine, Trp (W) tryptophan, Tyr (Y) tyrosine, Val (V) valine.
- 2) **Nucleic Acids and Relatives:** adenosine 5'-mono-, di-, or triphosphate (AMP, ADP, or ATP), cyclic adenosine 3', 5'-monophosphate (cAMP), complementary deoxyribonucleic acid (cDNA) (= copy DNA), complementary ribonucleic acid (cRNA) (=copy RNA), cytidine 5'-mono-, di-, or triphosphate (CMP, CDP, or CTP), deoxyadenosine 5'-mono-, di-, or triphosphate (dAMP, dADP, or dATP), deoxycytidine 5'-mono-, di-, or triphosphate (dCMP, dCDP, or dCTP), deoxyguanosine 5'-mono-, di-, or triphosphate (dGMP, dGDP, or dGTP), deoxynucleotide triphosphate (dNTP), deoxyribonucleic acid (DNA), double stranded ribonucleic acid (dsRNA), flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN), guanosine 5'-mono-, di-, or triphosphate (GMP, GDP, or GTP), genomic deoxyribonucleic acid (genomic DNA, gDNA), messenger ribonucleic acid (mRNA), micro ribonucleic acid (miRNA), nicotinamide adenine dinucleotide (NAD, NAD⁺), nicotinamide adenine dinucleotide phosphate (NADP, NADP⁺), nicotinamide adenine dinucleotide phosphate, reduced (NADPH), nicotinamide adenine dinucleotide, reduced (NADH), ribonucleic acid (RNA), ribosomal ribonucleic acid (rRNA), short interfering (silencing) ribonucleic acid (siRNA), transfer ribonucleic acid (tRNA), uridine 5'-mono-, di-, or triphosphate (UMP, UDP, or UTP).
- 3) **Buffers:** 2-[bis(2-hydroxyethyl)amino]ethanesulfonic acid (BES), *N,N'*-bis(2-hydroxyethyl)glycine (Bicine), 2-[bis(hydroxyethyl)amino]-2-(hydroxymethyl)-1-propane-1,3-diol (BisTris), 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonic acid (CHAPS), ethylenediaminetetraacetate (EDTA), ethyleneglycolbis-(β -aminoethyl ether)-*N,N,N',N'*-tetraacetic acid (EGTA), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 2-(*N*-morpholilino)ethanesulfonic acid (MES), 3-(*N*-morpholino)propanesulfonic acid (MOPS), 1,4-piperazinediethanesulfonic acid (PIPES), *N*-tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid (TES), *N*[tris(hydroxymethyl)methyl]glycine (Tricine), tris(hydroxymethyl)-aminomethane (Tris), tris-buffered saline (TBS).

- 4) **Related Terms:** base pair (bp), kilobase pair (kb), megabase pair (Mb), nucleotide (nt), Coenzyme A and its derivative (CoA), amplified fragment length polymorphism (AFLP), restriction fragment length polymorphism (RFLP), diethylaminoethyl (DEAE), bovine serum albumin (BSA), GTP-binding protein (G protein), gene of interest (= target gene) (GOI), quantification cycle (in PCR) (C_q).

For terms not present in the lists above, contributors should consult the recommendations of the International Union of Biochemistry and Molecular Biology (<http://www.chem.qmul.ac.uk/iubmb/>)

B. Abbreviations and Descriptions of Molecular Assays

- 1) **Gel Electrophoresis Assays:** polyacrylamide gel electrophoresis (PAGE), sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE), Southern blotting (gel electrophoresis with blotting and detection of DNA, capitalized because it is named after E.M. Southern = DNA blotting), northern blotting (gel electrophoresis with blotting and detection of RNA, not capitalized = RNA blotting), western blotting (gel electrophoresis with blotting and detection of proteins, not capitalized = protein blotting), two-dimensional polyacrylamide gel electrophoresis (2D-electrophoresis, 2-DE, or 2D-PAGE), electrophoretic mobility shift assay (EMSA).
- 2) **Polymerase Chain Reaction Assays:** polymerase chain reaction (PCR), rapid amplification of cDNA ends (RACE), real-time polymerase chain reaction (real-time PCR or qPCR), reverse transcription polymerase chain reaction (RT-PCR), RNA interference (RNAi).
- 3) **Microarray Assays:** expressed sequence tag (EST), genome survey sequence (GSS), high-throughput genome sequence (HTGS), sequence-tagged site (STS), tentative unique genes (TUGs).
- 4) **Protein/Peptide Assays:** dithiothreitol (DTT), electrospray ionization mass spectrometry (ESI-MS), enzymelinked immunosorbent assay (ELISA), fast-protein liquid chromatography (FPLC), high-performance liquid chromatography (HPLC), immunoglobulin G, M, etc. (IgG, IgM, etc.), isoelectric point (pI), liquid chromatography tandem mass spectrometry (LC MSMS), matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS), nuclear magnetic resonance (NMR), radioimmunoassay (RIA), tryptic digestion.

III. Presentation of Molecular Biological Content

A. PCR and PCR Primers

In the **Methods and Materials** include sequences of all forward and reverse PCR primers in one table and subdivide the table by the function of the primers. Functions might include generating probes for a northern blot experiment; primers for real-time PCR analysis; genomic or cDNA isolation; 3'-or 5'-RACE; or tissue localization assays. Examples of these tables can be found in: Lu et al. 2007, p. 1360; Xiu and Dong 2007, p. 950; Saltzmann et al. 2008, p. 1404; Li et al. 2008, p. 1594; Xiu et al. 2008, p. 489; Shivaji et al. 2010, p. 182. Primers should be written from 5' to 3', without any space between the abbreviations for the bases.

Provide the quantity of nucleic acid template used in the reaction; the quantities (molar concentrations) of all reagents (primers, dNTPs, and the polymerase); and the reaction volume. Also provide cycling conditions including temperatures and times, number of cycles, as well as the brand and model of the thermocycler and the method of isolation of the PCR products. Optimally, isolated PCR products (or products that have been cloned into plasmids) should be sequenced in both senses (5' to 3' and 3' to 5'), however, sequencing one strand of the product at least twice to confirm the identity of each base is also acceptable. The commercial entity or center responsible for the sequencing should also be indicated in the manuscript.

B. Real-time PCR (qPCR)

For guidelines to proper experimental design, calculation of gene expression, and statistical analysis in real-time PCR experiments, contributors should refer to Bustin et al. (2009) and Rieu and Powers (2009). Bustin et al. (2009) provide a checklist (in their Table 1) of all of the information necessary to consider and include when describing a high-quality real-time PCR experiment for publication. For the Journal, background data on real-time PCR should be included in the supplemental data section. Because real-time PCR is a quantification method highly dependent on the quality of the template (RNA or cDNA) and the efficiency of the amplification (C_q), authors should include tables that show the efficiencies of the amplification reaction for each gene monitored, including the reference gene(s). An example of this can be found in Zhang et al. (2009) (Supplemental Data Tables 1 and 2). R-squared values from the standard curves for genes analyzed by real-time PCR (including the reference gene(s) and the no-treatment controls) should also be included in these tables.

Authors are strongly encouraged to characterize the expression of multiple (typically three) reference genes for these experiments. Authors should determine the stability of the expression of their reference genes under their experimental conditions (validation of the reference genes) and then compare the normalized data from their treatments and provide statistical information on those comparisons that they view as most relevant. The expression of the reference gene should not differ significantly among the tissues, life stages, castes, etc. that are

under investigation (based on the same experimental procedures used for the genes of interest). This cannot necessarily be predicted *a priori*, so multiple reference genes should be checked and then the gene that best fits the criterion of "no significant change" should be used as the reference.

C. Genomics

Papers that use genomic, proteomic, or metabolomic approaches to study problems relevant to chemical ecology are appropriate for the Journal, provided that they are focused on chemical ecological themes (Tittiger 2004; Liu et al. 2007; McLean et al. 2007; Lawrence et al. 2008; Leiss et al. 2009). Manuscripts describing studies that utilize these approaches should go beyond cataloging and provide unique insights into regulatory networks, mechanisms, or demonstrations of ecological function.

For guidelines on genomics content, the Journal follows the conventions adopted by the journal *Molecular Endocrinology* and those of the Functional Genomics Data Society (<http://www.mged.org/>). Authors submitting expression or tiling microarray datasets must clearly identify in the **Methods and Materials** the platform, which includes the name of the vendor or array source, the name of the genechip or array and its version (e.g., Affymetrix Murine Genome U74v2 Set). Authors must thoroughly describe the filtering criteria used to evaluate the raw data and provide references for the statistical methods used to analyze the data. Filtered gene lists provided as supplemental data must be provided as Excel spreadsheets and not PDFs. Specific examples of supplemental data accompanying gene expression profiling manuscripts can be found in Ohlsson Teague et al. (2009); supplemental data available at: (<http://mend.endojournals.org/cgi/content/full/me.2008-0387/DC1>).

Experiments must be described according to minimum information about a microarray experiment (MIAME) guidelines (Brazma et al. 2001). Papers based on unreplicated gene expression profiling experiments will not be accepted for publication. Supporting experiments (i.e., realtime PCR or RT-PCR) are not a substitute for the biological replications of a genomics gene expression profiling experiment. However, microarray data must be verified by using an independent RNA sample; not with the RNAs used in the microarray experiments. The Journal recommends consultation with a statistician regarding experimental design and data analysis strategies for all gene expression profiling experiments.

Upon acceptance of a paper, authors of microarray datasets are required to have submitted the complete dataset to the Gene Expression Omnibus (GEO) (<http://www.ncbi.nlm.nih.gov/geo/>), which is the public gene expression archive of National Institute of Biotechnology Information (NCBI); or to ArrayExpress (<http://www.ebi.ac.uk/microarray-as/ae/>), which is the corresponding database of the European Bioinformatics Institute (EBI). Submission requirements can be found at either website. Datasets can be submitted confidentially to NCBI GEO prior to publication and held there until the paper is in press or published.

D. Proteomics

For guidelines on proteomics content, the Journal generally follows the conventions adopted by the journal *Molecular and Cellular Proteomics* (<http://www.mcponline.org/site/misc/itoea.xhtml> or <http://www.mcponline.org/site/misc/PhialdelphiaGuidelinesFINALDRAFT.pdf>).

When presenting protein identifications based on a single peptide sequence, authors should provide precursor mass, charge, and mass error (Carletti et al., 2008, see Table 1 on p. 809). Papers that use proteomic methods and report molecular structures, whether based on x-ray crystallography, NMR, or computational modeling, will be accepted only after the structural coordinates have been deposited in the Worldwide Protein Data Bank: <http://rcsb-deposit.rutgers.edu/> and <http://pdbdep.protein.osaka-u.ac.jp/>. Whenever possible, peptide sequences should be identified with matching accession numbers (see below).

E. GenBank and other Database Accessions for Manuscripts Reporting New Amino Acid or Nucleotide Sequences

Here, the Journal generally follows the conventions and publication guidelines adopted by *Insect Biochemistry and Molecular Biology* (http://www.elsevier.com/wps/find/journaldescription.cws_home/390/authorinstructions#). The GenBank sequence database is an open access, annotated collection of all publicly available nucleotide sequences and their protein translations. This database is produced at NCBI at the U.S. National Library of Medicine as part of the International Nucleotide Sequence Database Collaboration. The data are organized by accession numbers, which are unique identifiers in bioinformatics allocated to nucleotide and protein sequences to allow tracking of different versions of that sequence record and the associated sequence in GenBank or the Worldwide Protein Data Bank.

The Journal requires that authors deposit any novel nucleic acid sequences described in their paper in the GenBank database (<http://www.ncbi.nlm.nih.gov/genbank/> or <http://www.ncbi.nlm.nih.gov/WebSub/?tool=genbank>) or in the EMBL data library (<http://www.ebi.ac.uk/embl/>) and report the accession number in the **Methods and Materials** of the manuscript. Deposition of such data should be made at the time of submission. Papers will be accepted only after these data have been deposited in one of these databases. Note that the information provided in the GenBank accessions (nucleotide sequence, deduced protein sequence, etc.) is extensive and often makes the presentation of a DNA sequence and its protein translation unnecessary as a figure in the manuscript.

New amino acid sequences should be deposited in the Worldwide Protein Data Bank: <http://rcsb-deposit.rutgers.edu/> and <http://pdbdep.protein.osaka-u.ac.jp/>.

There are different types of accession numbers in use based on the type of sequence cited, each of which uses a different code. Authors should explicitly mention the *type of accession number together with the actual number*, bearing in mind that an error in a letter or number can result in a dead link in the online version of the article. Please use the following format: accession number type ID: xxxx (e.g., MMDB ID: 12345; PDB ID: 1TUP).

References

- BRAZMA, A., HINGAMP, P., QUACKENBUSH, J., SHERLOCK, G., SPELLMAN, P., STOECKERT, C., AACH, J., ANSORGE, W., BALL, C. A., CAUSTON, H. C., GAASTERLAND, T., GLENISSON, P., HOLSTEGE, F. C., KIM, I. F., MARKOWITZ, V., MATESE, J. C., PARKINSON, H., ROBINSON, A., SARKANS, U., SCHULZEKREMER, S., STEWART, J., TAYLOR, R., VILO, J., and VINGRON, M. 2001. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nature Genetics* 29:365–371.
- BUSTIN, S.A., BENES, V., GARSON, J.A., HELLEMANS, J., HUGGETT, J., KUBISTA, M., MUELLER, R., NOLAN, T., PFAFFL, M. W., SHIPLEY, G. L., VANDESOMPELE, J., and WITTEWER, C. T. 2009. The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clinical Chem.* 55:611–622.
- CARLETTI, P., MASI, A., SPOLAORE, B., POLVERINO de LAURETO, P., DE ZORZI, M., TURETTA, L., FERRETTI, M., and NARDI, S. 2008. Protein expression changes in maize roots in response to humic substances. *J. Chem. Ecol.* 34:804–818.
- DA COSTA, G., LAMY, E., CAPELA E SILVA, F., ANDERSEN, J., SALES BAPTISTA, and COELHO, A. V. 2008. Salivary amylase induction by tannin-enriched diets as a possible countermeasure against tannins. *J. Chem. Ecol.* 34: 376–387.
- HAN, Y., WANG, Y., BI, J. –L., YANG, X. –Q., HUANG, Y., ZHAO, X., HU, Y., and CAI, Q. –N. 2009. Constitutive and induced activities of defense-related enzymes in aphid-resistant and aphidsusceptible cultivars of wheat. *J. Chem.Ecol.* 35:176–182.
- LAWRENCE, S. D., NOVAK, N. G., JU, C. J. –T., and COOKE, J. E. K. 2008. Potato, *Solanum tuberosum*, defense against Colorado potato beetle, *Leptinotarsadecemlineata* (Say): Microarray gene expression profiling of potato by Colorado potato beetle regurgitant treatment of wounded leaves. *J. Chem. Ecol.* 35:1013–1025.
- LEISS, K. A., CHOI, Y. H., ABDEL-FARID, I. B., VERPOORTE, R., and KLINKHAMER, P. G. L. 2009. NMR Metabolomics of thrips (*Frankliniella occidentalis*) resistance in *Senecio* hybrids. *J. Chem. Ecol.* 35:219–229.
- LI, H. –L., ZHANG, Y. –L., GAO, Q. –K., CHENG, J. –A., and LOU, B.–G. 2008. Molecular identification of cDNA, immunolocalization and expression of a putative odorant-binding protein from an Asian honey bee, *Apis cerana cerana*. *J. Chem. Ecol.* 34:1593–1601.
- LIU, X., BAI, J., HUANG, L., ZHU, L., LIU, X. WENG, N., REESE, J. C., HARRIS, M., STUART, J. J., and CHEN, M. –S. 2007. Gene expression of different wheat genotypes during attack by virulent and avirulent Hessian fly (*Mayetiola destructor*) larvae. *J. Chem. Ecol.* 33:2171–2194.
- LU, D., LI, X., LIU, X., and ZHANG, Q. 2007. Identification and molecular cloning of putative odorant-binding proteins and chemosensory protein from the Bethyloid wasp, *Scleroderma guani* Xiao et Wu. *J. Chem. Ecol.* 33:1259–1375.
- MCLEAN, L., HURST, J. L., GASKELL, C. J., LEWIS, J. C. M., and BEYNON, R. J. 2007. Characterization of cauxin in the urine of domestic and big cats. *J. Chem. Ecol.* 33:1997–2009.
- OHLSSON TEAGUE, E. M. C., VAN DER HOEK, K. H., VAN DER HOEK, M. B., PERRY, N., WAGAARACHCHI, P., ROBERTSON, S. A. PRINT, C. G., and HULL, L. M. 2009. MicroRNA-regulated pathways associated with endometriosis. *Molec. Endocrinol.* 23:265–275.
- PELLETIER, J., GUIDOLIN, A., SYED, Z., CORNEL, A. J., and LEAL, W. S. 2010. Knockdown of a mosquito odorant-binding protein involved in the sensitive detection of oviposition attractants. *J. Chem. Ecol.* 36:245–246.
- RIEU, I. and POWERS, S. J. 2009. Real-time quantitative RT-PCR: Design, calculations, and statistics. *The Plant Cell* 21:1031–1033. doi:[10.1105/tpc.109.066001](https://doi.org/10.1105/tpc.109.066001)
- ROBERTSON, D. H. L., HURST, J. L., SEARLE, J. B., GÜNDÜZ, İ., and BEYNON, R. J. 2007. Characterization and comparison of major urinary proteins from the house mouse, *Mus musculus domesticus*, and the aboriginal

mouse, *Mus macedonicus*. *J. Chem. Ecol.* 33:613–630.

SALTZMANN, K. D., GIOVANINI, M. P., ZHANG, C., and WILLIAMS, C. E. 2008. Virulent Hessian fly larvae manipulate the free amino acid content of host wheat plants. *J. Chem. Ecol.* 34:1401–1410.

SETHI, A., MCAUSLANE, H. J., RATHINASABAPATHI, B., NUSSLY, G. S., and NAGATA, R. T. 2009. Enzyme induction as a possible mechanism for latex-mediated insect resistance in Romaine lettuce. *J. Chem. Ecol.* 35:190–200.

SEYBOLD, S.J. 2004. The eighth day of discovery: Molecular biology comes to chemical ecology. Editorial Preface to Special Molecular Issue of *The Journal of Chemical Ecology*. *J. Chem. Ecol.* 30:2327–2333.

SHIVAJI, R., CAMAS, A., ANKALA, A., ENGELBERTH, J., TUMLINSON, J. H., WILLIAMS, W. P., WILKINSON, J. R., and LUTHE, D. S. 2010. Plants on constant alert: Elevated levels of jasmonic acid and jasmonate-induced transcripts in caterpillar-resistant maize. *J. Chem. Ecol.* 36:179–191.

TITTIGER, C. 2004. Functional genomics and insect chemical ecology. *J. Chem. Ecol.* 30:2335–2358.

WEN, Z., ZENG, R. S., NIU, G., BERENBAUM, M. R., and SCHULER, M. A. 2009. Ecological significance of induction of broad-substrate cytochrome P450s by natural and synthetic inducers in *Helicoverpa zea*. *J. Chem. Ecol.* 35:183–189.

XIU, W. –M. and DONG, S. –L. 2007. Molecular characterization of two pheromone-binding proteins and quantitative analysis of their expression in the beet armyworm, *Spodoptera exigua* Hübner. *J. Chem. Ecol.* 33:947–961.

XIU, W. –M., ZHOU, Y. –Z., and DONG, S. –L. 2008. Molecular characterization and expression pattern of two pheromonebinding proteins from *Spodoptera littura* (Fabricius). *J. Chem. Ecol.* 34:487–498.

ZHANG, Z. –C., WANG, M. –W., LU Y. –B., and ZHANG, G. 2009. Molecular characterization and expression pattern of two general odorant binding proteins from the diamondback moth, *Plutella xylostella*. *J. Chem. Ecol.* 35:1188–1196.