

regulate photosynthetic capacity in green fruit through their regulation of chlorophyll accumulation and chloroplast development and ultimately contribute to sugars that accumulate in ripe fruit.

As in many other plants, two *GLK* genes are present and expressed in tomato, but in fruit, *SlGLK2* mRNA predominates and accumulates in a spatial pattern consistent with chlorophyll biosynthesis and chloroplast development. All *u/u* cultivars examined contain a *Slglk2* allele encoding a truncated loss-of-function GLK protein. Our results suggest that breeding selections for the *u* fruit trait that is helpful for harvesting methods may have had an unintended negative impact on fruit quality because suboptimal chloroplasts develop, and consequently, ripe fruit sugar and lycopene levels decrease. Manipulation of *GLK* levels or spatial expression patterns represents an opportunity to recover and enhance production and quality traits in tomato and other crop species.

References and Notes

1. L. Butler, *J. Hered.* **43**, 25 (1952).
2. A. F. Yeager, *Proc. Am. Soc. Hort. Sci.* **33**, 512 (1935).
3. S. M. Kinzer, S. J. Schwager, M. A. Mutschler, *Theor. Appl. Genet.* **79**, 489 (1990).
4. C. M. Rick, *Hilgardia* **42**, 493 (1974).
5. S. D. Tanksley, J. Hewitt, *Theor. Appl. Genet.* **75**, 811 (1988).
6. G. A. Kemp, I. L. Nonnecke, *Can. J. Plant Sci.* **40**, 306 (1960).
7. C. M. Rick, L. Butler, *Adv. Genet. Incorpor. Mol. Gen. Med.* **8**, 267 (1956).
8. M. T. Waters, J. A. Langdale, *EMBO J.* **28**, 2861 (2009).
9. S. Hetherington, R. Smillie, W. Davies, *J. Exp. Bot.* **49**, 1173 (1998).
10. M. M. Blanke, F. Lenz, *Plant Cell Environ.* **12**, 31 (1989).
11. S. Carrara, A. Pardossi, G. F. Soldatini, F. Tognoni, L. Guidi, *Photosynthetica* **39**, 75 (2001).
12. A. J. Matas *et al.*, *Plant Cell* **23**, 3893 (2011).
13. T. Manzara, P. Carrasco, W. Gruissem, *Plant Mol. Biol.* **21**, 69 (1993).
14. M. Sugita, W. Gruissem, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7104 (1987).
15. L. A. Wanner, W. Gruissem, *Plant Cell* **3**, 1289 (1991).
16. B. Piechulla, W. Gruissem, *EMBO J.* **6**, 3593 (1987).
17. B. Piechulla, R. E. Glick, H. Bahl, A. Melis, W. Gruissem, *Plant Physiol.* **84**, 911 (1987).
18. D. W. Fitter, D. J. Martin, M. J. Copley, R. W. Scotland, J. A. Langdale, *Plant J.* **31**, 713 (2002).
19. J. A. Langdale, *Plant Cell* **23**, 3879 (2011).
20. M. T. Waters *et al.*, *Plant Cell* **21**, 1109 (2009).
21. M. T. Waters, E. C. Moylan, J. A. Langdale, *Plant J.* **56**, 432 (2008).
22. H. Nakamura *et al.*, *Plant Cell Physiol.* **50**, 1933 (2009).
23. A. Bravo-García, Y. Yasumura, J. A. Langdale, *New Phytol.* **183**, 133 (2009).
24. Y. Yasumura, E. C. Moylan, J. A. Langdale, *Plant Cell* **17**, 1894 (2005).
25. A. H. Paterson *et al.*, *Nature* **335**, 721 (1988).
26. S. D. Tanksley, M. A. Mutschler, C. M. Rick, in *Genetic Maps*, S. J. O'Brien, Ed. (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1987), pp. 655–669.
27. Y. Eshed, D. Zamir, *Genetics* **141**, 1147 (1995).
28. Materials and methods are available as supplementary materials on Science Online.
29. J. Rohrmann *et al.*, *Plant J.* **68**, 999 (2011).
30. E. M. A. Enfissi *et al.*, *Plant Cell* **22**, 1190 (2010).
31. Y. Liu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 9897 (2004).

Acknowledgments: Minimum Information About a Microarray Experiment (MIAME)—compliant microarray data are available at <http://tedd.bti.cornell.edu>, and at <http://www.ebi.ac.uk/arrayexpress> (accession E-MEXP-3652). F. Carrari and A. Fernie provided *S. pennellii* *SlGLK2*, and J. Maloof provided *S. habrochaites* *SlGLK2* sequences. The U.S. Department of Agriculture (USDA)/National Institute of Food and Agriculture Solanaceae Coordinated Agricultural Project provided potato data. We are grateful to the Tomato Genome Consortium and the SOL Genomics Network for prepublication access to the tomato genome sequence. The *S. pennellii* introgression lines were provided by the C. M. Rick Tomato Genetics Resource Center; the *S. pimpinellifolium* populations were provided by the Instituto de Hortofruticultura Subtropical y Mediterranea "La Mayora," Consejo Superior de Investigaciones Científicas; and both populations are available by request from the sources. The *AtGLK*-expressing lines were provided by Mendel Biotechnology and Seminis/Monsanto Vegetable Seeds. *SlGLK2*, the corresponding lines, and the F2 10-1 IL x M82 population lines and seeds are available from J.J.G. without restriction. Seminis/Monsanto will make available, upon request, and under a material transfer agreement indicating they are to be used for noncommercial purposes, the following lines: LexA:AtGLK1:p35S:LexA-Gal4; LexA:AtGLK1:pLTP:LexA-Gal4; LexA:AtGLK1:pRbcS:LexA-Gal4; LexA:AtGLK1:pPDS:LexA-Gal4; LexA:AtGLK2:p35S:LexA-Gal4; LexA:AtGLK2:pLTP:LexA-Gal4; LexA:AtGLK2:pRbcS:LexA-Gal4;

LexA:AtGLK2:pPDS:LexA-Gal4; plus the T63 control line. Other biological materials are available by request from A.L.T.P. or J.J.G. A.L.T.P., T.H., K.L.-C., R.F.-B., and A.B.B. have filed a provisional U.S. patent application UC #2011-841, "Introduction of wild species *GLK* genes for improved ripe tomato fruit quality," through the University of California. A.L.T.P. and A.B.B. have filed the U.S. patent application #2010/0154078, "Transcription factors that enhance traits in plant organs," through Mendel Biotechnology. Assistance from B. Blanco-Ulate, S. Phothiset, S. Reyes, A. Abraham, L. Gilani, and G. Arellano is gratefully acknowledged. J. Langdale provided helpful advice regarding *GLK* phylogeny and nomenclature. G. Adamson and P. Kysar, Electron Microscopy (EM) Laboratory, University of California Davis Medical Center did the EM work. University of California Discovery and partners funded the pepper analysis and the initial investigations of the *Arabidopsis* *GLKs*. The Vietnam Education Foundation supported C.N. Fundación Genoma España ESPOL Project provided partial funding to A.G. USDA—Agricultural Research Service, USDA—National Research Initiative (2007-02773), and NSF (Plant Genome Program IOS-0923312) provided support to J.J.G.

Supplementary Materials

www.sciencemag.org/cgi/content/full/336/6089/1711/DC1
Materials and Methods
Figs. S1 to S8
Tables S1 to S8
References (32–58)

21 March 2012; accepted 5 June 2012
10.1126/science.1222218

The Paleozoic Origin of Enzymatic Lignin Decomposition Reconstructed from 31 Fungal Genomes

Dimitrios Floudas,¹ Manfred Binder,¹ Robert Riley,² Kerrie Barry,² Robert A. Blanchette,³ Bernard Henrissat,⁴ Angel T. Martínez,⁵ Robert Otillar,² Joseph W. Spatafora,⁶ Jagjit S. Yadav,⁷ Andrea Aerts,² Isabelle Benoit,^{8,9} Alex Boyd,⁶ Alexis Carlson,¹ Alex Copeland,² Pedro M. Coutinho,⁴ Ronald P. de Vries,^{8,9} Patricia Ferreira,¹⁰ Keisha Findley,¹¹ Brian Foster,² Jill Gaskell,¹² Dylan Glotzer,¹ Paweł Górecki,¹³ Joseph Heitman,¹¹ Cedar Hesse,⁶ Chiaki Hori,¹⁴ Kiyohiko Igarashi,¹⁴ Joel A. Jurgens,³ Nathan Kallen,¹ Phil Kersten,¹² Annegret Kohler,¹⁵ Ursula Kües,¹⁶ T. K. Arun Kumar,¹⁷ Alan Kuo,² Kurt LaButti,² Luis F. Larrondo,¹⁸ Erika Lindquist,² Albee Ling,¹ Vincent Lombard,⁴ Susan Lucas,² Taina Lundell,¹⁹ Rachael Martin,¹ David J. McLaughlin,¹⁷ Ingo Morgenstern,²⁰ Emanuelle Morin,¹⁵ Claude Murat,¹⁵ Laszlo G. Nagy,¹ Matt Nolan,² Robin A. Ohm,² Aleksandrina Patyshakuliyeva,⁹ Antonis Rokas,²¹ Francisco J. Ruiz-Dueñas,⁵ Grzegorz Sabat,²² Asaf Salamov,² Masahiro Samejima,¹⁴ Jeremy Schmutz,²³ Jason C. Slot,²¹ Franz St. John,¹² Jan Stenlid,²⁴ Hui Sun,² Sheng Sun,¹¹ Khajamohiddin Syed,⁷ Adrian Tsang,²⁰ Ad Wiebenga,⁹ Darcy Young,¹ Antonio Pisabarro,²⁵ Daniel C. Eastwood,²⁶ Francis Martin,¹⁵ Dan Cullen,¹² Igor V. Grigoriev,^{2*} David S. Hibbett^{1*}

Wood is a major pool of organic carbon that is highly resistant to decay, owing largely to the presence of lignin. The only organisms capable of substantial lignin decay are white rot fungi in the Agaricomycetes, which also contains non-lignin-degrading brown rot and ectomycorrhizal species. Comparative analyses of 31 fungal genomes (12 generated for this study) suggest that lignin-degrading peroxidases expanded in the lineage leading to the ancestor of the Agaricomycetes, which is reconstructed as a white rot species, and then contracted in parallel lineages leading to brown rot and mycorrhizal species. Molecular clock analyses suggest that the origin of lignin degradation might have coincided with the sharp decrease in the rate of organic carbon burial around the end of the Carboniferous period.

Lignin is a heterogeneous polymer that provides strength and rigidity to wood, protects cellulose and hemicellulose from microbial attack, and is the major precursor of

coal (*1*). Genomic studies of wood decay organisms have focused on model fungal systems for white rot (in which all plant cell wall components are degraded), such as *Phanerochaete*

chrysosporium (2), and brown rot (in which lignin is modified but not appreciably degraded), such as *Postia placenta* (3) and *Serpula lacrymans* (4). However, these species represent just two of the 18 recognized orders of Agaricomycetes, of which five contain brown rot taxa. To reconstruct the evolution of lignin decay mechanisms, we analyzed 31 diverse fungal genomes, including 12 newly sequenced species of Agaricomycotina (Table 1). The new genomes comprise six white rot species, five brown rot species, and one mycoparasite, representing nine orders (Fig. 1 and figs. S1 to S5) (5).

To estimate phylogenetic relationships among these taxa, we constructed data sets using 71 or 26 single-copy genes, with varying alignment criteria and treatments for fast-evolving sites, yielding matrices of 10,002 to 34,257 amino acids, which we analyzed with maximum likelihood (ML) and Bayesian methods (5–7). All but six nodes receive maximal support values in all analyses, and the rest are strongly supported (bootstrap $\geq 99\%$

or posterior probability ≥ 0.99) in at least three analyses. The tree topology is consistent with prior analyses and resolves four independent brown rot lineages (Fig. 1A and fig. S6).

We searched all 31 genomes for 27 gene families encoding oxidoreductases and carbohydrate-active enzymes (CAZymes) that have been implicated in wood decay (Table 1). CAZymes, particularly those acting on crystalline cellulose, are abundant in white rot genomes, which have 61 to 148 (average 87) copies of genes encoding CAZymes, representing 14 to 17 gene families, whereas brown rot genomes have 32 to 68 copies (average 46) from 9 to 12 families. The ectomycorrhizal (ECM) *Laccaria bicolor* resembles brown rot species in this regard, possessing 28 CAZyme genes in eight families (Table 1). Notably, glycoside hydrolase (GH) families GH6 and GH7, which include cellobiohydrolases that are involved in the attack of crystalline cellulose (8), are present in all white rot lineages, but they are absent in brown rot lineages (except Boletales) and *L. bicolor*. Similar patterns of enrichment in white rot genomes are shown by genes encoding GH61 enzymes, which have a copper-dependent oxidative mechanism for disrupting crystalline cellulose (9), and cellulose binding modules (CBM1), which effectively increase the concentration of the enzymes on the surface of crystalline cellulose (10) (Table 1).

To gain access to cellulose, wood-decaying fungi must overcome or circumvent lignin; thus, we focus on fungal class II peroxidases (PODs), which degrade lignin in *P. chrysosporium* and other species (11) (figs. S7 to S19). We classified PODs into four major groups, including three ligninolytic forms—lignin peroxidase (LiP), manganese peroxidase (MnP), and versatile peroxidase (VP)—and a fourth POD type, defined here as “generic peroxidase” (GP), which is expected to include nonligninolytic low-redox potential peroxidases with catalytic properties similar to those of the peroxidase of *Coprinopsis cinerea* or the product of the *nopA* gene in *P. chrysosporium* (5, 12). LiPs possess a tryptophan residue on the surface of the enzyme corresponding to Trp¹⁷¹ in *P. chrysosporium* LiP-H8 that enables direct oxidation of lignin compounds via long-range electron transfer; MnPs possess two or three residues corresponding to Glu³⁵, Glu³⁹, and Asp¹⁷⁵ of *P. chrysosporium* MnP1 that function in binding Mn (13). VPs possess both the Trp¹⁷¹ homolog and Mn-binding residues, whereas all are lacking in GPs.

Consistent with a central role for PODs in lignin degradation, white rot species have 5 to 26 copies (average 14) of genes encoding ligninolytic PODs, but all brown rot species lack these enzymes, as do the ECM *L. bicolor*, the soil saprotroph *C. cinerea*, and *Schizophyllum commune*, which has been regarded as a white rot fungus but has a limited capacity to degrade lignin (14). Moreover, analyses of gene diversification with binary state speciation analysis (15) confirmed

that the rate of duplication of POD genes is elevated in white rot lineages versus non-white rot lineages (5).

To reconstruct the functional evolution of PODs, we performed Bayesian and ML analyses (6, 16) using the GPs of Ascomycota as outgroups, and we estimated the ancestral states of the key residues of ligninolytic PODs using BayesTraits (17). Our results indicate that the ancestor of all PODs likely lacked the Mn-binding and Trp¹⁷¹ residues of MnP, LiP, and VP, suggesting that it was nonligninolytic (Fig. 1B). The most recent ancestor of all ligninolytic Agaricomycete PODs is reconstructed as an MnP, suggesting that there was a single origin of LiP (gain of Trp¹⁷¹ and loss of Mn-binding residues), with parallel expansions in the *P. chrysosporium* and *Trametes versicolor* (Polyporales, each with 10 LiP copies; Fig. 1B and figs. S7 and S17). We also identified two origins of VP in the Polyporales, where *T. versicolor* and *Dichomitium squalens* each have three VP copies (Fig. 1B and fig. S7). VPs are also produced in the “oyster mushroom” *Pleurotus ostreatus* (Agaricales) (18), indicating further convergent evolution of this class of enzymes.

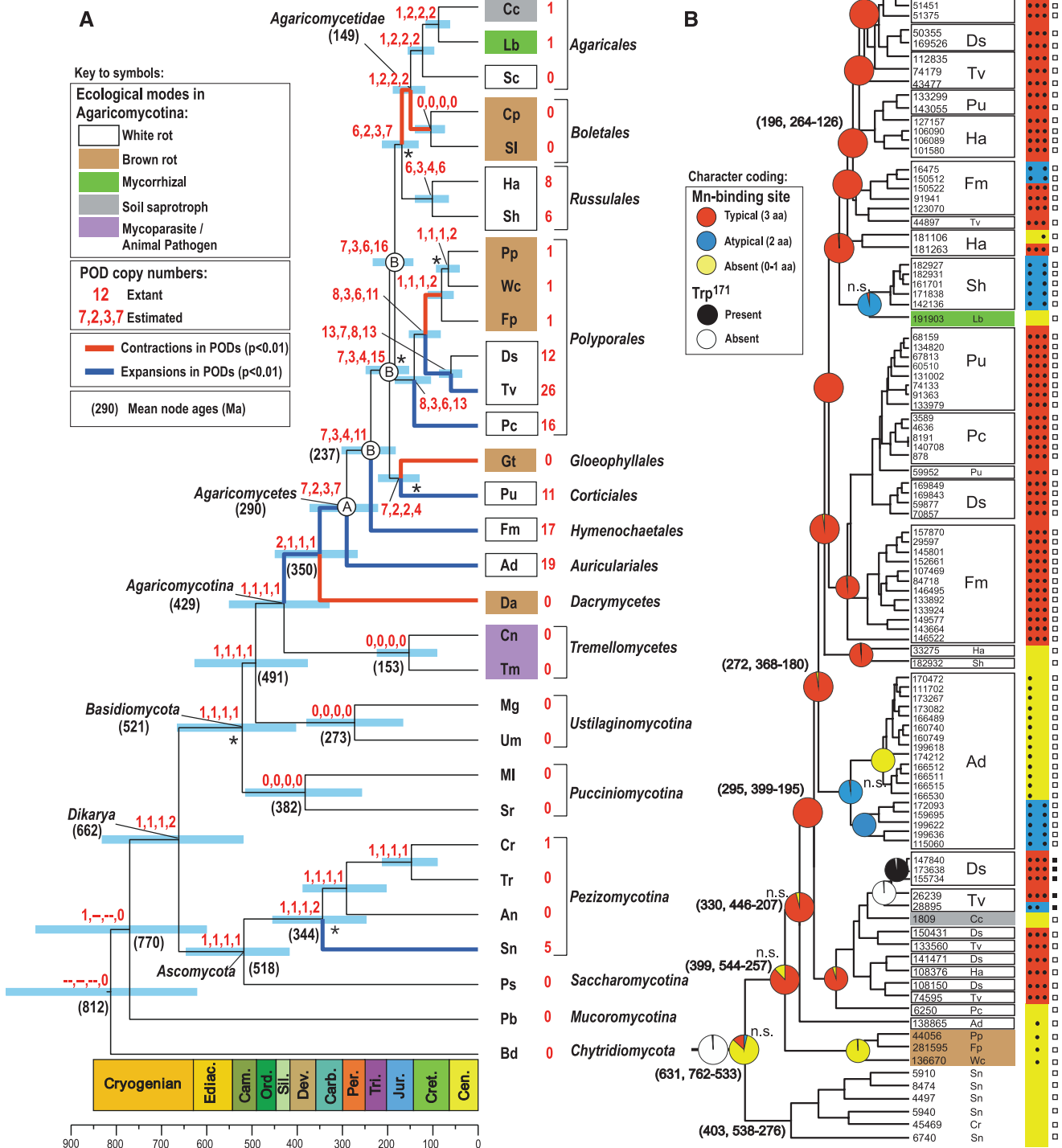
To localize the diversification of PODs in the organismal phylogeny, we performed gene tree/species tree reconciliation analyses using CAFE (19), Notung (20), and DrML (21). All methods suggest that a single POD gene copy was present in the common ancestor of Basidiomycota, with parallel losses in lineages leading to the Pucciniomycotina, Ustilaginomycotina, Tremellomycetes, and *Dacryopinax* sp. (Fig. 1A). Diversification of PODs occurred in the lineage leading to the most recent common ancestor of the Agaricomycetes (node “A” in Fig. 1A), which is reconstructed as having two to seven POD gene copies in the various analyses. In addition, reconciliation analyses suggest that the ancestor of the Agaricomycetes possessed one or two genes encoding dye-decolorizing peroxidases (DyP), which are heme peroxidases that have been shown to degrade lignin model compounds (22), as well five to eight genes encoding oxidases (including glyoxal oxidase) involved in peroxide generation (5, 23). Collectively, these results suggest that the ancestor of Agaricomycetes was a white rot species that possessed a ligninolytic system with PODs, DyPs, and multiple pathways for H₂O₂ production.

The “backbone” nodes in the Agaricomycete phylogeny (labeled “B” in Fig. 1A) are reconstructed as having 3 to 16 POD gene copies, which suggests that the white rot mechanism was retained throughout the early evolution of Agaricomycetes. Subsequently, all reconciliation analyses suggest that there were parallel expansions of POD genes in terminal lineages, leading to white rot species in five orders (Auriculariales, Hymenochaetales, Corticiales, Russulales, and Polyporales). In contrast, parallel contractions of PODs are resolved within lineages leading to the brown rot *Dacryopinax* sp., *Gloeophyllum trabeum*, the Boletales, and the brown rot Polyporales,

¹Biology Department, Clark University, Worcester, MA 01610, USA. ²U.S. Department of Energy Joint Genome Institute, Walnut Creek, CA 94598, USA. ³Department of Plant Pathology, University of Minnesota, St. Paul, MN 55108, USA. ⁴Architecture et Fonction des Macromolécules Biologiques, Aix-Marseille Université, CNRS UMR 6098, 13288 Marseille Cedex 9, France. ⁵Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, E-28040 Madrid, Spain. ⁶Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR 97331, USA. ⁷Environmental Genetics and Molecular Toxicology Division, Department of Environmental Health, University of Cincinnati College of Medicine, Cincinnati, OH 45267, USA. ⁸Microbiology and Kluwer Centre for Genomics of Industrial Fermentation, Utrecht University, Padualaan 8, 3584 CH Utrecht, Netherlands. ⁹CBS-KNAW Fungal Biodiversity Centre, Uppsalalaan 8, 3584 CT Utrecht, Netherlands. ¹⁰Department of Biochemistry and Molecular and Cellular Biology and Institute of Biocomputation and Physics of Complex Systems, Zaragoza University, 50009, Zaragoza, Spain. ¹¹Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710, USA. ¹²USDA Forest Products Laboratory, Madison, WI 53726, USA. ¹³Institute of Informatics, Warsaw University, Warsaw, 02-097, Poland. ¹⁴Department of Biomaterial Sciences, Graduate School of Agricultural and Life Sciences, University of Tokyo, 1-1-1, Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan. ¹⁵INRA, UMR 1136, INRA-Nancy Université, Interactions Arbres/Microorganismes, 54280 Champenoux, France. ¹⁶Molecular Wood Biotechnology and Technical Mycology, Büsgen-Institute, Georg-August-University Göttingen, Büsgenweg 2, D-37077 Göttingen, Germany. ¹⁷Department of Plant Biology, University of Minnesota, St. Paul, MN 55108, USA. ¹⁸Departamento de Genética Molecular y Microbiología, Facultad de Ciencias Biológicas, Pontificia Universidad Católica de Chile, Casilla 114-D, Santiago 833-1010, Chile. ¹⁹Department of Applied Chemistry and Microbiology, Viikki Biocenter, P.O. Box 56, Biocenter 1, University of Helsinki, FIN-00014 Helsinki, Finland. ²⁰Centre for Structural and Functional Genomics, Concordia University, 7141 Sherbrooke Street West, Montreal, Quebec H4B 1R6, Canada. ²¹Department of Biological Sciences, Vanderbilt University, Nashville, TN 37235, USA. ²²University of Wisconsin Biotechnology Center, Madison, WI 53726, USA. ²³HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA. ²⁴Department of Forest Mycology and Pathology, Swedish University of Agricultural Sciences, Box 7026, Ulls v 26A, 750 07 Uppsala, Sweden. ²⁵Genetics and Microbiology Research Group, Public University of Navarre, 31006 Pamplona, Spain. ²⁶College of Science, University of Swansea, Singleton Park, Swansea SA2 8PP, UK.

*To whom correspondence should be addressed. E-mail: ivgrigoriev@lbl.gov (I.V.G.); dhibbett@clarku.edu (D.S.H.)

Fig. 1. (A) Organismal phylogeny (chronogram) produced with BEAST from a 26-gene data set. Light blue bars are 95% highest posterior density intervals for node ages; mean ages of selected nodes (millions of years) are in parentheses. Blue and red branches indicate significant expansion and contraction, respectively, of PODs inferred using CAFE. Numbers in red at node names are POD gene counts. Numbers in red at nodes, separated by commas, are numbers of POD gene copies estimated with CAFE, Notung (with two different edge weight threshold settings), and DrML, respectively. The node labeled A is the ancestor of Agaricomycetes; nodes labeled B are “backbone” nodes in Agaricomycetes (see text). Asterisks indicate nodes that do not receive maximal support in all analyses (see fig. S6 for support values). See Table 1 for full species names. **(B)** POD gene phylogeny estimated in BEAST with ancestral state reconstructions for manganese-binding site (colored pies) and Trp¹⁷¹ residues (black and white pies) estimated with BayesTraits. Bars to right of gene IDs indicate presence of functional residues (13). Mean ages for selected nodes in parentheses are followed by 95% highest posterior density ranges.



end of the Permo-Carboniferous was caused, at least in part, by the evolution of lignin decay capabilities in white rot Agaricomycetes.

References and Notes

- J. M. Robinson, *Geology* **18**, 607 (1990).
- D. Martínez et al., *Nat. Biotechnol.* **22**, 695 (2004).
- D. Martínez et al., *Proc. Natl. Acad. Sci. U.S.A.* **106**, 1954 (2009).
- D. C. Eastwood et al., *Science* **333**, 762 (2011).
- See supplementary materials on Science Online.
- A. Stamatakis, *Bioinformatics* **22**, 2688 (2006).
- N. Lartillot, T. Lepage, S. Blanquart, *Bioinformatics* **25**, 2286 (2009).
- P. Baldrian, V. Valášková, *FEMS Microbiol. Rev.* **32**, 501 (2008).
- R. J. Quinlan et al., *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15079 (2011).
- D. Guillén, S. Sánchez, R. Rodríguez-Sanoja, *Appl. Microbiol. Biotechnol.* **85**, 1241 (2010).
- A. T. Martínez, F. J. Ruiz-Dueñas, M. J. Martínez, J. C. Del Río, A. Gutiérrez, *Curr. Opin. Biotechnol.* **20**, 348 (2009).
- L. Larrondo, A. González, T. Perez Acle, D. Cullen, R. Vicuña, *Biophys. Chem.* **116**, 167 (2005).
- F. J. Ruiz-Dueñas et al., *J. Exp. Bot.* **60**, 441 (2009).
- O. Schmidt, W. Liese, *Holzforchung* **34**, 67 (1980).
- R. G. FitzJohn, W. P. Maddison, S. P. Otto, *Syst. Biol.* **58**, 595 (2009).
- A. J. Drummond, A. Rambaut, *BMC Evol. Biol.* **7**, 214 (2007).
- D. Barker, A. Meade, M. Pagel, *Bioinformatics* **23**, 14 (2007).
- F. J. Ruiz-Dueñas, E. Fernández, M. J. Martínez, A. T. Martínez, *C. R. Biol.* **334**, 795 (2011).
- T. De Bie, N. Cristianini, J. P. Demuth, M. W. Hahn, *Bioinformatics* **22**, 1269 (2006).
- D. Durand, B. V. Halldórsson, B. Vernet, *J. Comput. Biol.* **13**, 320 (2006).
- P. Górecki, G. J. Burleigh, O. Eulenstein, *BMC Bioinformatics* **12** (suppl. 1), S15 (2011).
- M. Hofrichter, R. Ullrich, M. J. Pecyna, C. Liers, T. Lundell, *Appl. Microbiol. Biotechnol.* **87**, 871 (2010).
- A. Vanden Wymelenberg et al., *Appl. Environ. Microbiol.* **72**, 4871 (2006).
- M. Krings, N. Dotzler, G. Galtier, T. N. Taylor, *Mycoscience* **52**, 18 (2011).
- S. P. Stubblefield, T. N. Taylor, *Bot. Gaz.* **147**, 116 (1986).
- R. A. Berner, *The Phanerozoic Carbon Cycle: CO₂ and O₂* (Oxford Univ. Press, Oxford, 2004).

Acknowledgments: The work conducted by the U.S. Department of Energy Joint Genome Institute was supported by the Office of Science of the U.S. Department of Energy under contract

DE-AC02-05CH11231. Also supported by the Assembling the Fungal Tree of Life (AFTOL) project under NSF awards DEB-0732968 (D.S.H.), DEB-0732993 (J.W.S.), and DEB-0732550 (D.J.M.). We thank R. H. Petersen for the strain of *A. delicata*. The organisms *A. delicata* and *Dacryopinax* sp. were obtained in Costa Rica and can only be used for research purposes. Assemblies and annotations of the 12 genomes reported here are available from the JGI fungal portal MycoCosm (<http://jgi.doe.gov/fungi>) and from DDBJ/EMBL/GenBank under the following accessions: AFV000000000, AEIT000000000, AEUS000000000, AEID000000000, AEJ000000000, AEHC000000000, AFV000000000, AEGM000000000, AEGX000000000, AEJ000000000, AFV000000000, and AEHD000000000. Aligned sequence data for organismal and gene family phylogenies and molecular clock analyses, and secretome results are available at DRYAD (<http://dx.doi.org/10.5061/dryad.5k3t47p0>).

Supplementary Materials

www.sciencemag.org/cgi/content/full/336/6089/1715/DC1

Materials and Methods

Supplementary Text

Tables S1 to S16

Figs. S1 to S22

References

12 March 2012; accepted 7 May 2012

10.1126/science.1221748

Leucine-tRNA Initiates at CUG Start Codons for Protein Synthesis and Presentation by MHC Class I

Shelley R. Starck,¹ Vivian Jiang,¹ Mariana Pavon-Eternod,² Sharanya Prasad,¹ Brian McCarthy,³ Tao Pan,² Nilabh Shastri^{1*}

Effective immune surveillance by cytotoxic T cells requires newly synthesized polypeptides for presentation by major histocompatibility complex (MHC) class I molecules. These polypeptides are produced not only from conventional AUG-initiated, but also from cryptic non-AUG-initiated, reading frames by distinct translational mechanisms. Biochemical analysis of ribosomal initiation complexes at CUG versus AUG initiation codons revealed that cells use an elongator leucine-bound transfer RNA (Leu-tRNA) to initiate translation at cryptic CUG start codons. CUG/Leu-tRNA initiation was independent of the canonical initiator tRNA (AUG/Met-tRNA_i^{Met}) pathway but required expression of eukaryotic initiation factor 2A. Thus, a tRNA-based translation initiation mechanism allows non-AUG-initiated protein synthesis and supplies peptides for presentation by MHC class I molecules.

In almost all nucleated cells, newly translated polypeptides supply antigenic precursors for loading major histocompatibility complex (MHC) class I molecules (1). Peptide-loaded MHC class I (pMHC I) molecules reveal the presence of viral or mutated proteins to circulating cytotoxic T cells (CTLs), which bind pMHC I through their T cell receptors to eliminate infected or transformed cells. Antigenic precursors are translated from conventional AUG-initiated open reading frames (ORFs) by the canonical initiator transfer RNA (tRNA), Met-tRNA_i^{Met} (2, 3). Cryptic, non-AUG-initiated ORFs

(4–7) also generate pMHC I during viral infections (8–12) and oncogenesis (13–15) by unknown mechanisms. Cryptic CUG start codons can also be decoded with leucine at the initiation stage of translation in mammalian cells (5–7, 16). However, this decoding is incompatible with the current model of translation, which indicates that ribosomes are preloaded with initiator Met-tRNA_i^{Met} before recognition of AUG or even non-AUG start codons (17).

Our previous study suggested that translation of antigenic precursors from a CUG start codon using leucine represents a distinct initiation pathway (16). To determine the molecular mechanism of CUG/leucine initiation, we first screened a series of compounds described as inhibitors of eukaryotic protein synthesis (18) using primer extension inhibition analysis (toeprinting) (19) of AUG-YL8 and CUG-YL8 mRNA ribosome initiation complexes (fig. S1). We found that NSC119893, which inhibits Met-tRNA_i^{Met}

association with eukaryotic initiation factor 2 (eIF2) (20), selectively inhibits AUG-YL8 initiation (table S1) in a dose-dependent fashion, whereas CUG-YL8 toeprints were resistant to NSC119893 treatment (Fig. 1A). Structurally unrelated protein synthesis inhibitors, such as suramine and aurin tricarboxylic acid, also inhibited AUG initiation yet enhanced initiation at the CUG start codon (table S1). In contrast, the small molecule acriflavine inhibited CUG initiation more than AUG initiation in a dose-dependent fashion (Fig. 1A). Thus, a structurally diverse set of compounds can distinguish ribosomal recognition of AUG and CUG start codons.

We next assessed the effect of these protein synthesis inhibitors on translation of antigenic precursors in living cells by biochemically analyzing peptides from extracts of cells transfected with the AUG-YL8 or CUG-YL8 plasmids (5). As expected, a single peak of antigenic activity—corresponding to the methionine-initiated peptide (MYL8)—was detected from AUG-YL8-transfected cells (Fig. 1B). Yet, CUG-YL8-transfected cells yielded the leucine-initiated peptide (LYL8) as well as the MYL8 peptide, arising from Met-tRNA_i^{Met} “wobble” initiation (Fig. 1C). Although NSC119893 inhibited the expression of MYL8 from AUG and CUG start codons, it did not inhibit decoding of the CUG initiation codon with leucine (LYL8) (Fig. 1, B to D). In contrast, and consistent with the toeprint analysis (Fig. 1A), translation of LYL8 was inhibited by acriflavine, a nucleic acid intercalator, whereas MYL8 initiation from either AUG or CUG codons was unaffected (Fig. 1, B to D). This inhibitor effect was not limited to translation of antigenic precursors, because NSC119893 inhibited AUG-GFP, but not CUG-GFP, expression (fig. S2, A and B). Conversely, CUG-GFP, but not AUG-GFP, expression was inhibited by acriflavine treatment in cultured cells

¹Division of Immunology and Pathogenesis, Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA. ²Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60637, USA. ³DNA Sequencing Facility, University of California, Berkeley, CA 94720, USA.

*To whom correspondence should be addressed. E-mail: nshastri@berkeley.edu

forces that generate and maintain gene clusters for specialized plant metabolism (12), such clusters, once identified, are invaluable for decoding of complex, taxonomically restricted pathways in plants. One need only compare the pace of dissecting noscapine synthesis to the nearly two decades required to fully decipher the pathway for morphine to fully appreciate the power of such metabolic Rosetta stones. Low-cost, high-throughput sequencing has increasingly driven biosynthetic gene cluster identification in plants;

as our ability to apply molecular genetic tools (such as virus-induced gene silencing) expands to more nonmodel plant species with unique biochemistries, we can anticipate that even more dark recesses of specialized plant metabolism will be illuminated.

References

1. T. Winzer *et al.*, *Science* **336**, 1704 (2012); 10.1126/science.1220757.
2. G. M. Cragg, D. J. Newman, *J. Ethnopharmacol.* **100**, 72 (2005).
3. D. K. Liscombe, S. E. O'Connor, *Phytochemistry* **72**, 1969 (2011).

4. M. Mahmoudian, P. Rahimi-Moghaddam, *Anti-Cancer Drug Discov.* **4**, 92 (2009).
5. M. Frey *et al.*, *Science* **277**, 696 (1997).
6. A. M. Takos *et al.*, *Plant J.* **68**, 273 (2011).
7. S. Swaminathan *et al.*, *Plant Cell* **21**, 3315 (2009).
8. B. Field *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 16116 (2011).
9. B. Field, A. E. Osbourn, *Science* **320**, 543 (2008).
10. K. Shimura *et al.*, *J. Biol. Chem.* **282**, 34013 (2007).
11. X. Qi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 8233 (2004).
12. H. Y. Chu, E. Wegel, A. Osbourn, *Plant J.* **66**, 66 (2011).

10.1126/science.1225473

EVOLUTION

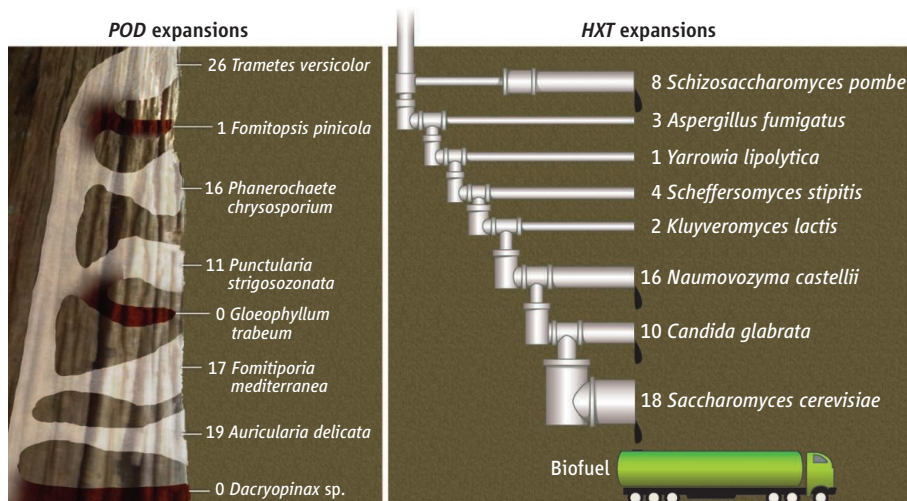
Endless Rots Most Beautiful

Chris Todd Hittinger

Fungal rots that decay wood were not prominent among the “endless forms most beautiful” that Darwin chronicled, but if he had known of the biochemical and evolutionary processes at work, they might have been. Woody plants fix an extraordinary amount of carbon during their lifetimes, building towering trees of decay-resistant lignocellulose. On page 1715 of this issue, Floudas *et al.* performed deep phylogenetic sampling of fungal genomes to describe how white rot Agaricomycetes fungi have evolved an arsenal of enzymes to degrade lignin and unlock its stored carbon (1).

As plants invaded land, lignin provided the rigidity necessary for vascular plants to grow above their rivals and move water and nutrients over long distances (2). Lignin is a dizzying web of polymerized phenylalanine derivatives with dozens of combinations of modifications and cross-links that make wood structurally sound and render it inaccessible to all but the most persistent chemical and biological assaults (2–5). The success of lignin-rich plants in the swamp forests of the Carboniferous Period created many of the coal-rich deposits that fueled the Industrial Revolution. But little carbon is buried today, in part because of white rot fungi. Floudas *et al.*'s analysis of 12 newly sequenced species of Agaricomycotina provides a treasure trove of wood-decaying enzymes to test and industrialize (5, 6) as well as remarkable insights into the genomics of adaptive shifts, gene duplication and diversification, and parallel evolution.

Laboratory of Genetics, Wisconsin Bioenergy Initiative, Great Lakes Bioenergy Research Center, University of Wisconsin, Madison, WI 53706, USA. E-mail: cthittinger@wisc.edu



Parallel gene family expansions in fungi. (Left) Approximate evolutionary relationships (left to right, 300 million years ago to the present) and numbers of *POD* genes (also represented by width) in white and brown rot basidiomycetes. (Right) Relationships and numbers of *HXT* genes in ascomycetes; ethanol drops indicate taxa that ferment aerobically, including *S. cerevisiae*, the main producer of liquid biofuels.

Broad taxonomic and ecological sampling of rots allowed Floudas *et al.* to use state-of-the-art phylogenomic and ancestral state inference to trace the history of the Agaricomycetes and their myriad lignocellulolytic enzymes. Fifteen gene families deviate from the null expectation of a random birth-death process of gene duplication and loss, and instead exhibit significant lineage-specific expansions and contractions. For example, glycoside hydrolases, multicopper oxidases, and dye-decolorizing peroxidases expanded during the evolution of the lignin-degrading life-style of white rots, while contractions occurred in brown rot lineages that do not appreciably degrade lignin.

The most striking expansions and contractions are apparent in the fungal class II

peroxidases (PODs), which are the primary lignin-degrading enzymes in white rots (7). The common ancestor of Agaricomycetes is inferred to have been a white rot with a modest repertoire of manganese peroxidase PODs. Indeed, nearly all modern white rot fungi possess several manganese peroxidases, whereas only a few species possess lignin peroxidases and versatile peroxidases capable of directly oxidizing aromatic rings. The branching pattern of the Agaricomycetes suggests that PODs independently expanded along multiple white rot lineages through gene duplication and independently contracted to zero or one *POD* genes in at least three brown rot lineages (see the figure). The few PODs found in non-white rot fungi all lack key sequences associated with lignolytic

activity (4, 8). Thus, the aggressive decay of lignin by white rots requires a diverse, refined suite of PODs and other enzymes.

The *POD* gene family expansions and contractions provide clear examples of genome content evolving in concert with the white or brown rot niches, but the parallel changes observed in some of the most lignolytic PODs are even more conspicuous. In both the lignin peroxidases (non-manganese-binding) and the versatile peroxidases (manganese-binding), natural selection repeatedly found the same solution to direct oxidation of lignin by changing a key external residue to tryptophan, enabling long-range electron transfer (4, 9). Parallel alterations of the manganese-binding residues were also observed in several PODs, although it is unclear whether these are losses of function or trade-offs that enable new enzymatic properties. Adjusting specific residues is reminiscent of the spectral tuning that has occurred repeatedly in opsins to enable animals to see various light wavelengths (10).

As the white rots became experts in degrading the complex carbon stored in

wood by duplicating and diversifying their *POD* gene families, *Saccharomyces* yeasts mastered fermenting simple sugars in sap and fruit, in part by duplicating and diversifying their hexose transporters (HXTs; see the figure). Most yeasts (and filamentous ascomycetes) have one to five *HXT* genes, but species that have evolved to ferment glucose in the presence of oxygen have extensively duplicated and diversified this gene family (11). The champion fermenter *S. cerevisiae* has 18 *HXT* genes that encode various specificities, capacities, and fine-tuned differential regulation (12, 13). *HXT* expansions occurred independently in *Schizosaccharomyces pombe*, another yeast that ferments aerobically. The secondary or promiscuous activities exhibited by PODs and HXTs probably facilitated specialization after gene duplication (14, 15).

The gene family expansions of white rots and fermentative yeasts complement both their ecologies and their potential uses in biofuel production. Current lignocellulosic biorefinery designs call for chemical

and enzymatic deconstruction of biomass by enzymes modified from white rots and other decay specialists, followed by conversion of the hexoses and pentoses to ethanol by yeast (6). Although these complex genomic systems took hundreds of millions of years to evolve, we are beginning to understand them only just in time to exploit them to meet our energy needs.

References

1. D. Floudas, *Science* **336**, 1715 (2012).
2. J. K. Weng, C. Chapple, *New Phytol.* **187**, 273 (2010).
3. W. Boerjan *et al.*, *Annu. Rev. Plant Biol.* **54**, 519 (2003).
4. K. E. Hammel, D. Cullen, *Curr. Opin. Plant Biol.* **11**, 349 (2008).
5. A. T. Martínez *et al.*, *Curr. Opin. Biotechnol.* **20**, 348 (2009).
6. M. E. Himmel *et al.*, *Science* **315**, 804 (2007).
7. M. Tien, C. P. Tu, *Nature* **326**, 520 (1987).
8. L. F. Larrondo *et al.*, *Biophys. Chem.* **116**, 167 (2005).
9. W. A. Doyle *et al.*, *Biochemistry* **37**, 15097 (1998).
10. S. Yokoyama, *Prog. Retin. Eye Res.* **19**, 385 (2000).
11. Z. Lin, W.-H. Li, *Mol. Biol. Evol.* **28**, 131 (2011).
12. S. Ozcan, M. Johnston, *Microbiol. Mol. Biol. Rev.* **63**, 554 (1999).
13. R. Wieczorke *et al.*, *FEBS Lett.* **464**, 123 (1999).
14. R. A. Jensen, *Annu. Rev. Microbiol.* **30**, 409 (1976).
15. J. Piatigorsky, G. Wistow, *Science* **252**, 1078 (1991).

10.1126/science.1224682

CHEMISTRY

Rethinking Chemical Reactions at Hyperthermal Energies

Xueming Yang,^{1,2} Timothy K. Minton,^{1,3} Dong Hui Zhang¹

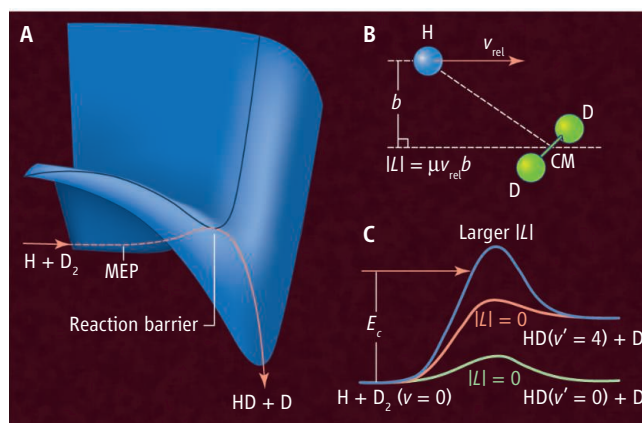
A chemical reaction usually involves the breaking of old chemical bonds and the formation of new bonds. Internal vibrational energy or translational energy of the reacting molecules could affect the dynamics of chemical bond breaking and formation processes in a major way. The effects of vibrational excitation on chemical reactivity have been extensively investigated for decades (1, 2), especially for highly vibrationally excited molecules. Studies of the translational effects on chemical dynamics have mainly focused on energies near the reaction barriers; reaction dynamics at hyperthermal translational energies

(far above the barriers) has received much less attention. On page 1687 of this issue, Jankunas *et al.* (3) report that the dynamics of the simplest chemical reaction, $H + D_2 \rightarrow HD + D$, change in unexpected ways at hyperthermal translational energies and they provide a theoretical explanation for this unusual behavior.

At translational energies slightly above the reaction threshold, reaction trajec-

tries normally follow the minimum energy path (MEP) (see the figure, panel A). As the energy increases, deviations from the MEP are possible and often change the dynamics in predictable ways. For example, in $H + D_2 \rightarrow HD + D$, reactive collisions along the collinear (MEP) pathway cause the HD products to rebound in the backward direction relative to the initial direction of the H

¹State Key Laboratory of Molecular Reaction Dynamics, Dalian Institute of Chemical Physics, Chinese Academy of Sciences (CAS), Dalian 116023, China. ²Department of Chemical Physics, University of Science and Technology of China, Hefei 230026, China. ³Montana State University, Bozeman, MT 59717, USA. E-mail: xmyang@dicp.ac.cn (X.Y.); tminton@montana.edu (T.K.M.); zhangdh@dicp.ac.cn (D.H.Z.)



A matter of momentum. (A) The potential energy surface and the MEP are shown for the $H + D_2 \rightarrow HD + D$ reaction ($|L| = 0$). (B) Reaction geometry for the $H + D_2$ reaction with impact parameter b (CM is center of mass); the angular momentum of the collision is $L = \mu v_{rel} b$, where μ is the reduced mass of the collision partners and v_{rel} is the relative velocity. (C) Energy diagram along the reaction coordinate for the $H + D_2$ reaction with zero and larger L , where a larger L collision that has a high centrifugal barrier will inhibit the reaction at collision energy E_c below this barrier.