Scientific Research and Essays Vol. 6(21), pp. 4651-4656, 30 September, 2011
Available online at http://www.academicjournals.org/SRE
DOI: 10.5897/SRE11.1257
ISSN 1992-2248 ©2011 Academic Journals

Full Length Research Paper

Bibliographic and citation analysis of publications from a Thai Research Institute with a defined lifespan: The Institute of Molecular Biology and Genetics, 1999 to 2009

Anchalee Nirachanon and Duncan R. Smith*

Institute of Molecular Biosciences, Mahidol University, Salaya Campus, 25/25 Phuttamontol Sai 4, Salaya, Nakorn Pathom, 73170 Thailand.

Accepted 17 August, 2011

This study sought to explore the research output of a single research institute in Thailand with a defined life span (1999 to 2009) and to assess the impact of this research through the use of citation analysis. A total of 228 primary research articles met the inclusion criteria (indexed primary research publication) with a combined total of 1696 published pages. Average article length was 7.43 pages. The average impact factor (IF) of the journals in which the papers were published was 2.48 based upon the year of publication. A total of 212 of the 228 papers (93%) had been cited with a cumulative 2376 citations giving an average citation rate of 10.42 citations per paper. This study provides the first detailed assessment of research productivity and impact of the work of a Thai Research Institute and may have value in benchmarking research in institutes in other developing countries.

Key words: Citation, developing nation, h-factor, impact, research output.

INTRODUCTION

Science and technology research are viewed as an indispensible part of a countries economic (Hyung, 1988; Mansfield, 1991) and social (Hyung, 1988; Sitthi-Amorn and Somrongthong, 2000) development. However, developing countries where the requirement for development is possibly the greatest traditionally spend a smaller proportion of their budget on research and development (R&D) than do developed countries. While developing countries spend some 2 to 3% of their GNP on R&D (Sharif, 1986), developing countries seldom reach this level. For example Thailand previously spent approximately 0.25% of its GDP (range 0.24 to 0.26 between 2001 and 2006), but this has dropped recently to 0.21% in 2007 (Taharnklaew, 2010) a figure significantly lower than other nations (Liefner and Schiller, 2008). In comparison with other East Asian countries, Korea, Taiwan and Singapore spend more than 2% of GDP on R&D, China spends slightly over 1% while the Philippines spends less than 0.1% (Liefner and Schiller, 2008). The reasons for the low investment in R&D in developing countries is complex, but probably stems from a disconnect where those who formulate governmental policy and budget see R&D as "consumption" rather than a critical facet of long term economic development (Hyung, 1988). This is compounded by a failure of these countries to have a system to accurately judge the economic impact of research and development (Hyung, 1988). Even divorced from the long term economic considerations, methods to evaluate research are highly contentious. Overall methodologies for the assessment of research performance fall into two categories (Avital and Collopy, 2001), the "evaluative" which focuses a retrospective evaluation of scientific effort (papers, patents) and the "explanatory" which is a proactive effort to identify factors related to research output. Explanatory studies have suggested that age and experience are both associated

*Corresponding author. E-mail: duncan_r_smith@hotmail.com at. Tel: (662) 800 3624. Fax (662) 441-9906.

with output in that scientists become less productive as they age and that peak performance is achieved within the first 10 years of work (Avital and Collopy, 2001). Similarly, explanatory studies have defined a number of personality traits that are associated with academic excellence including dominance, aggression and independence (Avital and Collopy, 2001). However, of the two types of studies, evaluative studies often used to rank Institutions, programs and researchers themselves are the more common especially since the widespread accessibility of bibliographic databases (Falagas et al., 2008).

While at the national level, assessment of productivity (or output) is often undertaken in the light of assessment of "input" factors such as total R&D expenditure (Sharif, 1986), at the lower levels (University, Institute or researcher) evaluation of output is often undertaken without consideration of these factors, although conversely "output" of an institute (or researcher) may play a critical role in determining "input' such as allocation of funding or manpower (Leary, 1985). To evaluate research at the institutional or researcher level, significant focus is placed on output in the form of papers or patents (or other innovations), although this tends to undervalue the role of teaching, particularly of post-graduate students in the evaluation process (Im et al., 1998). Currently, there are a number of ways in which output can be measured including simple paper counts, citations, impact factor analysis (Avital and Collopy, 2001) and more recently h-index analysis (Hirsch, 2005). All of these can serve to bench mark both productivity and impact, although all of them have significant advantages and disadvantages as analytical tools. In Thailand, there have been few formal bibliographic analysis of research output, and those studies that have been undertaken usually consider research output at the national or University level (Ruenwongsa and Panijpan, 1995; Svasti and Asavisanu, 2006) and to date only one study has attempted to evaluate the actual impact of the research undertaken at this level (Sombatsompop et al., 2010). The Institute of Molecular Biology and Genetics, Mahidol University was established as an independent Institute in 1999 from a previously existing unit under the auspices of the Institute of Science and Technology for Research and Development, Mahidol University. After several years of growth, the two Institutes were re-merged to form the Institute of Molecular Biosciences, Mahidol University in 2009. In this respect, the defined lifespan of the Institute of Molecular Biology and Genetics (11 years inclusive) offers a unique opportunity to examine the output of a Thai research Institute over a clearly defined period and to assess the impact of that research.

MATERIALS AND METHODS

Data collection and analysis

Data on publications from the Institute of Molecular Biology and Genetics, Mahidol University Thailand was collected by examination of the Institutes own records, the Elsevier SciVerse Scopus database and the Institute for Scientific Information-Web of Science database (ISI-WOS). The inclusion criterion was at least one author whose affiliation address was the "Institute of Molecular Biology and Genetics, Mahidol University". While data on all publications was collected, further analysis was confined to primary research publications with the exclusion of reviews and conference proceedings. Individual paper citations were collected from both the Elsevier SciVerse Scopus database and the Institute for Scientific Information-Web of Science database (ISI-WOS) database with a reference date of 30/06/2011. Data on authorship and affiliation was collected from database searches and cross checked against original hard copies of the publications. The impact factor (IF) of the journal in which a paper was published was retrieved from the Institute for Scientific Information (ISI) Journal Citation Reports (JCR) for the year of publication with the exception that for papers published after 2009 (6 papers) the 2009 IF was used. Data analysis was undertaking using the Microsoft Excel spreadsheet program.

RESULTS

The Institute of Molecular Biology and Genetics, Mahidol University existed as an independent research Institute within Mahidol University between 1999 and 2009 (11 years inclusive). Analysis of the Institutes own records, the Elsevier SciVerse Scopus database and the Institute for Scientific Information-Web of Science database (ISI-WOS) identified a total of 244 publications with at least one author whose affiliation address was the "Institute of Molecular Biology and Genetics, Mahidol University". Four articles were reviews and three were identified as proceedings papers. Of the remaining 237 primary research papers, 228 had entries in Scopus while 223 had entries in ISI-WOS. Two of the papers listed in ISI-WOS were not found in the Scopus database reflecting the slightly different coverage of these two databases. The higher coverage of papers in the Scopus database is consistent with the larger number of journals covered by Scopus as compared to the ISI-WOS database (Falagas et al., 2008) and the Scopus database was therefore used as the primary data source in this study except where stated. Of the 228 primary research articles listed by Scopus, 222 were published between 1999 and 2009 inclusive, although several papers were published in 2010 (5 papers) or 2011 (1 paper) which reflected long journal acceptance to publication times. The 228 research papers were published in a total of 101 different journals, with an average of 2.25 papers published by the Institute in each of those 101 journals, with a range from 20 (Journal of Biochemistry and Molecular Biology) to 1 (59 different journals). Three journals besides the Journal of Biochemistry and Molecular Biology had more ten or more papers from the Institute, namely Biochemical and Biophysical Research Communications (11 papers), Biochemical Journal (11 papers) and Current Microbiology (10 papers).

The average number of research papers produced per year was nearly 21 (20.7; range 4 to 33), with the lowest

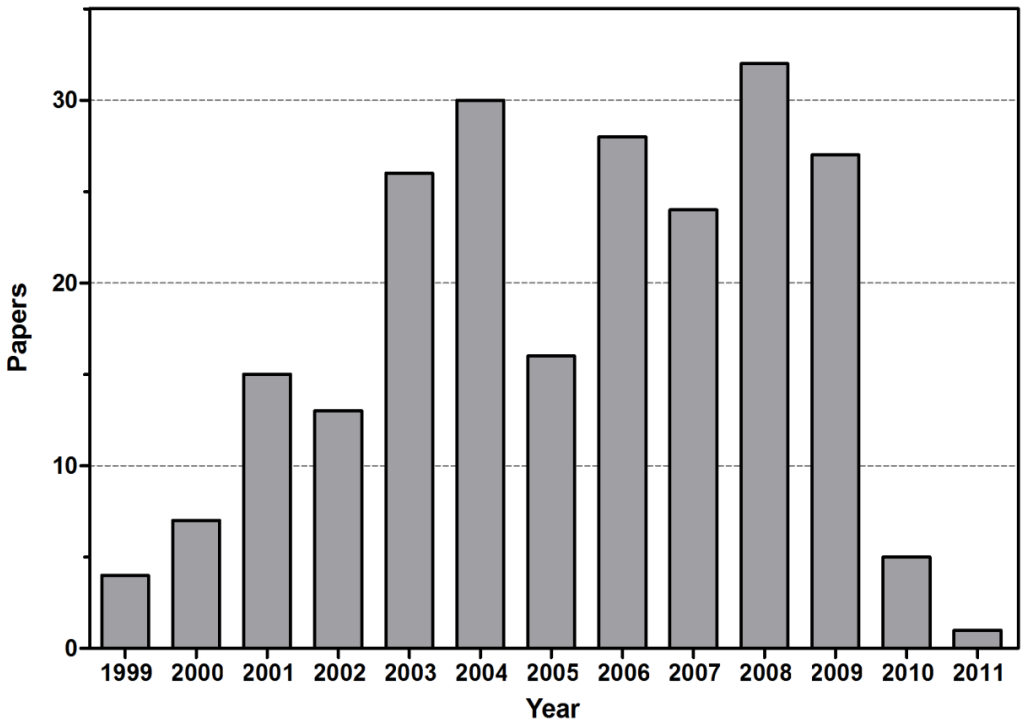

Figure 1. Research papers produced by year. Note that papers continued to be published with affiliation to the Institute after 2009.

numbers of papers being produced in the early years after formation of the Institute (Figure 1). The average published paper length was 7.43 pages (range 17 to 2 pages), with a cumulative total of 1696 pages or roughly 154 pages published per year. A total of 431 different authors were represented on the 228 research publications with the average number of authors per paper being 4.83 (range 2 to 14). The average number of publications per author was 2.5, although the range was extremely high (82 to 1). Although there was a great deal of overlap in authorship between papers, only 12 authors were present on 10 papers or more, and only 6 authors were present on more than 20 publications. Of the 228 papers, the corresponding author was affiliated with the institute in 66.2% of the publications (151/228). Of the remaining 77 papers (33.7%) with corresponding authors with affiliations other than the Institute, nearly 60% (46/77) of the authors were affiliated with a different Thai institution. In total less than 15% (31/228) of all of the research publications had a corresponding author from a non-Thai institution. Using the Scopus database, a total of 212 of the 228 papers (93%) have been cited with a cumulative 2376 citations (as of 30/06/2011). Surprisingly, only 16 papers have not been cited and half of these (8/16) were published in 2009/2010/2011. The average citations/per paper rate is 10.42 (range 88 to 0). The average yearly citation rate was 215.18/year with a range of 0 (1999) to 447 (2010) citations per year (Figure 2). When self citations for all authors were excluded, the total number of citations dropped to 1678, showing that nearly 30% of all citations were self citations (698/2376). This lowered the average citations per paper to 7.35. Using the ISI-WOS database 204 of the 223 paper had citation data (91.4%) with a cumulative citation total of 2155 and an average rate of citations per article of 9.66 (range 83 to 0).

Interestingly, both the Scopus database and the ISI-WOS database calculated an h index of 22 based upon the publications in their respective databases. Journal impact factors were available in the Institute for Scientific Information (ISI) Journal Citation Reports (JCR) for the year of publication for 209 of the publications. Some nineteen papers (8.3%) were published in journals for which there was no impact factor for the year in which the paper was published. Average journal impact factor was 2.48 with a range of 30.03 to 0.46.

DISCUSSION

Research efforts in both developing and developed nations tend to be centered around universities, and to a lesser extent autonomous or semi-autonomous research Institutes. In particular, research Universities, those with an equal or greater emphasis on research as opposed to teaching, serve as nuclei spearheading the research

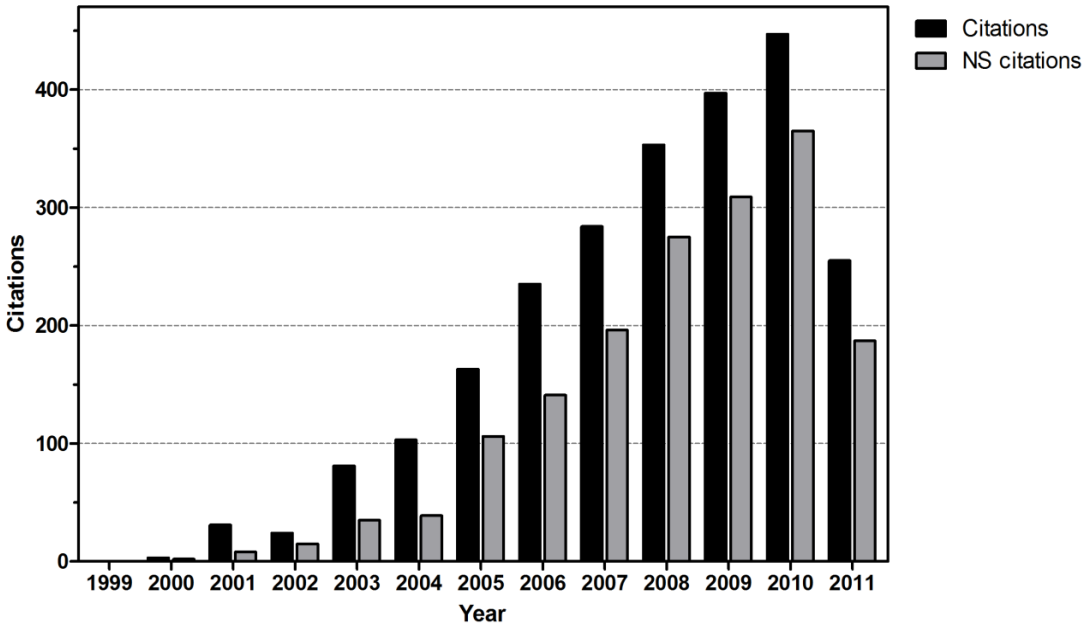

Figure 2. Citations per year on papers published. Both all citations and non-self (NS) citations are shown.

development of a country (Altback, 2011) and much of the available research budget in Thailand is spent funding research undertaken either directly at university departments or at research institutions associated with universities (Numprasertchai and Igel, 2005). The Institute of Molecular Biology and Genetics at Mahidol University, Thailand was created as an independent Institute within Mahidol University in 1999. At its maximum it consisted of less than 15 small groups predominantly consisting of one group leader and 3 to 6 Masters or PhD students undertaking research in a wide range of areas including agrobiology, protein structure and function and virology. During its lifespan, the Institute of Molecular Biology and Genetics published some 237 primary research articles, as well as a smaller number of reviews and conference proceedings. The small number of reviews was somewhat surprising and those reviews that were published all had a non-Thai first and last author, possibly implying a lack of confidence in compiling review articles because of perceived language difficulties. However, the average article length of the indexed 228 research papers was more than 7 pages, suggesting a greater confidence in the more structured environment of the standard research article. However, evaluation of research productivity generally centers on evaluation of the significance of the publications that arise as a result of the work and the simplest criterion of number of publications is seldom used as this offers no evaluation of the quality of the work produced and this metric is particularly susceptible to inflation by authors simply dividing their work into smaller publishable pieces.

The introduction in the 1960s of the Science Citation Index produced by the Institute for Scientific Information following the original concept of Garfield (1995) allowed the qualitative evaluation based upon the number of citations that papers have received. The citation number of a paper gives a reasonable idea of the broader impact of a paper, although it only details how many times a paper has been referenced by another paper, giving no value of whether the citing author agreed (positive context citation) or disagreed (negative context citation) with the cited paper. Garfield also proposed that analysis of citations could also lead to evaluation of the relative importance of scientific journals (Garfield, 1955) through calculations of an “impact factor” (IF). Garfield and Sher subsequently used impact factors as an aid to identifying journals for inclusion in the Science Citation Index (Garfield, 2006) and these were subsequently formally presented by the Institute for Scientific Information as the Journal Citation Reports (Garfield, 1972). The impact factor is calculated by dividing the total number of citations received by a particular journal in one year on articles published in the preceding two years by the total number of papers published over the preceding two years (Garfield, 1999). A high impact factor reflects a high average number of citations for papers in that journal, and the number can change every year as the impact factor is derived from actual citation and paper number data. However, IF does imply any value about the significance of a single work in that journal, but is merely reflective of the output of the journal as a whole (Avital and Collopy, 2001). For example a paper may be published in a high impact journal, but receive no citations. In this study, while the average impact of the

published papers was 2.48, there was an extremely broad range from more than 30 to approximately 0.46. Interestingly however, the highest impact paper only ranked 8th in terms of actual citations received to date. Direct citation counts therefore provide a better indicator of a papers actual impact in the field, albeit that those citations maybe negative context citations. In analysis of both researchers and research institutes, therefore, total citations may offer some indication of research impact, either as a discrete total or as an average value per paper. The most significant drawback to using direct citations as an evaluative tool is that citations can increase slowly over a long period of time and the most cited paper in this study obtained the highest number of citations in a year, some six years after the year publication (2004). However, overall, nearly 40% of all citations (936/2376) were within two years after the year of publication, suggesting that the work indeed has an immediate impact.

Including all citations, the average citations per paper was 10.42, a value that compares well with the average citations per article in life sciences of approximately 6 citations per article (Adler et al., 2009). More recently the concept of h index has gained increasing attention. This index was introduced by Hirsch who proposed that: “a scientist has index h if h of his/her Np papers and have at least h citations each, and the others (Np-h) have not more than h citations each” (Hirsch, 2005). While the index was originally proposed to assess individual scientists, more recently it has become applied to the research journals themselves, as well as to organizations. In this regard, the calculated h-index for the Institute studied here of 22 says that 22 papers (of the total 228 papers) had at least 22 citations. In this way, the h-index attempts to provide an indication of both quantity and impact, and it has been argued that h-index provides a better predictive indicator of scientific achievement than other biometric markers, including total citation count or citation per paper count (Hirsch, 2007). In the citation analysis, it was seen that some 698 out of the 2376 total citations (29.3%) were self citations. This figure, although seemingly high is consistent with other studies. For example, in a study of scientific production in Norway between 1981 and 1996, Aksnes found a self citation rate of 36% in a three year window period (Aksnes, 2003). While self citations are often seen as problematic (MacRoberts and MacRoberts, 1989; Seglen, 1997) in that they can serve to artificially inflate the apparent impact of the work, there are numerous legitimate reasons for self citation as part of the normal publishing process (Phelan, 1999). The problem of self citation can also be compounded by multiple authorships, and Aksnes (2003) found a strong positive correlation between self citation and the number of authors on a paper.

Self citation can be defined in a number of ways (Aksnes, 2003) including where a citation is considered a self citation if any author is present in common on the two papers. A more rigorous definition can include a citation is considered a self citation only if the first author is the same in the two studies (Garfield, 1979). In our analysis, even using the broad definition of self citation (that is, any author) there was still an average of 7.35 non-self citations per paper, a value that still compares well to the average of 6 citations per papers for life science articles (Adler et al., 2009). A surprisingly small proportion of papers (7%) had no citations and nearly 50% of these were within 2 years of publication. Evidence has shown that citation lives for papers are considerably longer than two years and although there is some difference between different disciplines, the citation life of a paper can be ten years or more (Adler et al., 2009). The figure of 7% uncited papers by the Institutes authors is exceptional given that studies have consistently shown that up to 50% of papers are never cited (Hamilton, 1990, 1991; Pendlebury, 1991) and that articles written by foreign authors generally exhibit a higher rate of uncitedness (Garfield, 1991). Analysis of publications for Thailand as a whole have shown that only 60% of articles are cited and that the average citation/article rate is 1.757 (Sombatsompop et al., 2010). However, Mahidol University is the most productive (in terms of publications) University in Thailand (Sombatsompop et al., 2010; Svasti and Asavisanu, 2006) and between 2007 and 2009 had a citation/article rate of between 2.5 and 2.8 (Sombatsompop et al., 2010). In light of this, the Institutes productivity in terms of publications and citations per article was creditable. However, the Institute had no undergraduate teaching commitment, unlike the majority of University faculties in Thailand and as such this should be factored in the final evaluation of the significance of the research undertaken.

ACKNOWLEDGEMENTS

DRS is supported by the Thailand Research Fund and Mahidol University.

REFERENCES

Adler R, Ewing J, Taylor P (2009). Citation statistics. Stat. Sci., 24: 1-14.

Aksnes DW (2003). A macro study of self citation. Scientometrics, 56: 235-246.

Altback PG (2011). The past, present, and future of the research university. Econ. Polit. Weekly, 46: 65-73.

Avital M, Collopy F (2001): Assessing research performance: Implications for selection and motivation. Case Western Reserve University, USA. Sprouts: Working papers on information systems, http://sprouts.aisnet.org/, 1(14): 1-14

Falagas ME, Pitsouni EI, Malietzis GA, Pappas G (2008). Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. FASEB J, 22: 338-342.

Garfield E (1955). Citation indexes for science; a new dimension in documentation through association of ideas. Science, 122: 108-11.

Garfield E (1972). Citation analysis as a tool in journal evaluation. Science, 178: 471-479.

Garfield E (1979). Is citation analysis a legitimate evaluation tool? Scientometrics, 1: 359-375.
Garfield E (1991). To Be An Uncited Scientist Is No Cause For Shame. The Scientist, 5: 12.
Garfield E (1999). Journal impact factor: a brief review. CMAJ, 161: 979-80.
Garfield E (2006). The history and meaning of the journal impact factor. JAMA, 295: 90-93.
Hamilton DP (1990). Publishing by--and for?--the numbers. Science, 250: 1331-1332.
Hamilton DP (1991). Research papers: who's uncited now? Science, 251: 25.
Hirsch JE (2005). An index to quantify an individual's scientific research output. Proc Natl Acad Sci U SA., 102: 16569-16572.
Hirsch JE (2007). Does the H index have predictive power? Proc. Natl. Acad. Sci USA., 104: 19193-19198.
Hyung SC (1988). Science policy mechanism and technology development strategy in the developing countries. Technol. ForecasT. Soc. Change., 33: 279-292.
Im KS, Kim KY, Kim JS (1998). A response to "Assessing research productivity: important but neglected considerations". Decision Line, 29: 12-15.
Leary RA (1985). A framework for assessing and rewarding a scientists research productivity. Scientometrics, 1-2: 29-38.
Liefner I, Schiller D (2008). Academic capabilities in developing countries-A conceptual framework with empirical illustrations from Thailand. Research Policy, 37: 276-293.
MacRoberts MH, MacRoberts BR. (1989). problems of citation analysis: A critical review. J. Am. Soc. Inf. Sci., 40: 342-349.
Mansfield E (1991). Academic research and industrial innovation. Res. Pol., 20: 1-12.
Numprasertchai S, Igel B (2005). Managing knowledge through collaboration: Multiple case studies of managing research in university laboratories in Thailand. Technovation, 25: 1173-1182.
Pendlebury DA. (1991). Science, citation, and funding. Science, 251: 1410-1411.
Phelan TJ (1999). A compendium of issues for citation analysis. Scientometrics, 45: 117-136.
Ruenwongsa P, Panijpan B (1995). Science and technology publications of State Universities in Thailand. J. Sci. Soc. Thailand., 21: 207-214.
Seglen PO (1997). Citations and journal impact factors: questionable indicators of research quality. Allergy, 52: 1050-1056.
Sharif MN (1986). Measurement of technology for national development. Technol. ForecasT. Soc. Change., 29: 119-172.
Sitthi-Amorn C, Somrongthong R (2000). Strengthening health research capacity in developing countries: a critical element for achieving health equity. BMJ, 321: 813-817.
Sombatsompop N, Markpin T, Ratchatahirun P, Yochai W, Wongkaew C, Premkamolnetr N (2010). Research performance evaluations of Thailand national research universities during 2007-2009. Inf. Dev., 26: 303-313.
Svasti J, Asavisanu R (2006). Update on Thai publications in ISI databases (1999-2005). ScienceAsia 32: 101-106.
Taharnklaew R (2010). The Scientist: Private Sector Power.